

Rethinking Mathematics Education

Alan H. Schoenfeld
University of California, Berkeley
Berkeley, CA USA 94720-1670
alans@berkeley.edu
+1 (510) 642-0968

Draft, November 26, 2021 + DOIs

Submitted for publication in:

Breaking Images: Iconoclastic Analyses of Mathematics and its Education

Edited by:

Brian Greer, Portland State University, Portland, OR, United States, brian1060ne@yahoo.com

Ole Skovsmose, Universidade Estadual Paulista, São Paulo, Brazil, osk@hum.aau.dk

David Kollosche, Universität Klagenfurt, Klagenfurt, Austria, david.kollosche@aau.at

A volume in the series “Studies on Mathematics Education and Society” from Open Book

Publishers: <https://www.openbookpublishers.com/section/125/1>

Rethinking Mathematics Education

Alan Schoenfeld
University of California, Berkeley

I had the good fortune to fall in love with mathematics as a child and to spend the early part of my career as a mathematician. Then, intrigued by Pólya's ideas about problem solving, I turned to mathematics education. The challenge as I understood it was, can we understand enough about mathematical thinking and problem solving to help students get good at it? Wouldn't it be great if increasing numbers of students could experience the power and beauty of mathematics, and even come to feel about it the way I do?¹

From my current vantage point the picture is far more complex. Over the years I have come to see mathematics as highly political, in the sense of realpolitik. It has become increasingly clear, as I have worked to help build a rigorous research base for productive change, that those of us who concern ourselves with the improvement of mathematics teaching and learning have far less to say about the enterprise than we might; that school math does little good in the "real world;" and that huge numbers of students are systematically excluded from participation in mathematics. Such realizations crystallized amidst the onset of Covid and the increase in racial tensions across the U.S. due to the murder of George Floyd. Schooling has been massively disrupted, with the concomitant exacerbation of already significant racial inequities. Yet nowhere have there been calls for re-thinking what is possible or appropriate – standards and testing remain unchallenged and meaningless concerns about "learning loss" predominate. Societal preoccupations and plausible academic goals conflict in uncomfortable ways.

This chapter provides a political/intellectual narrative, ultimately raising questions regarding the character of appropriate goals for mathematics education and how one might think about attaining them. It tells the larger, political story of my experiences as a researcher and developer as I have pursued deeper understandings of the nature of mathematical thinking, teaching and learning. The narrative takes a turn at the end, as I reflect on my mathematics-related experiences in recent years. I still love mathematics for its beauty and power, but I am deeply concerned about its non-use (except for those who have a professional need for it) in real-world contexts that matter. In my personal life I have made significant use of K-12 mathematics over the past few years – but in very different ways than the current goals of K-12 mathematics would suggest or support. It is time to rethink the rhetoric and reality of mathematics education. As I reflect on my real-world mathematical thinking during times of Covid, on the systematic exclusion of students from the mathematical pipeline, and on what I have found it important to emphasize in my ongoing problem solving courses, I think significant change is necessary. The question is, what should the goals of mathematics instruction be and how might we attain them? This chapter concludes with a proposal for change. That proposal is unlikely to gain traction, but perhaps something good can come from the issues it raises².

¹ Of course, what anyone takes pleasure in is a matter of taste. But we can imagine all students having opportunities to experience mathematics (or art, or sports, or literature) in ways that open up the potential for such pleasure.

² The issues that unfold in my narrative are sometimes grounded in the culture of the United States and sometimes general. Experiences within the sociopolitical context of the U.S. may or may not have analogs in other nations, but aspects of mathematical thinking are in large measure universal. My goals for mathematics instruction are thus a hybrid of the two.

Introduction

I am going to tell this tale as a chronological narrative, in that it reflects what might be called my sentimental education, from a young naïf to an older and perhaps somewhat wiser scholar-activist who has many of the same goals he had in his youth – to help students experience mathematics in ways that enable them to become mathematical sensemakers, experiencing its power and beauty – but is more cognizant both of the social obstacles that impede progress and of the failures of extant curricula.

The narrative begins with a focus on mathematical thinking, a brief recap of key ideas from my problem solving work. Once I understood what it is to be a powerful mathematical problem solver, I focused on understanding teaching, and then learning environments. Increasingly as well, I focused on things that can make a difference – curriculum, assessment, and professional development³. Second, there is my evolving political awareness. This was, of course, nascent early on; it was clear many decades ago that the nature of statewide standards and assessments shaped what was possible in the classroom. But the political nature of the standards process was not apparent to the young me – the “math wars” came after the first sets of standards were released. Similarly, the politics of professional development only became apparent to me when I engaged up close. These issues, in turn, paled in significance when society’s utter disregard for teachers became apparent during the onset of the Covid pandemic. Amid the chaos of 2020, there was good reason to re-think the purposes and impacts of education; but there was little work to do so. Increased income inequality and the blatancy of racial violence make the challenges we face that much more salient.

As has always been the case, my teaching and thinking are deeply intertwined; such issues make their way into my course on mathematical thinking and problem solving, and experiences in the course shape my thinking about what matters. It goes without saying that I write from a position of privilege, and that my history and my perspective have been shaped accordingly.

In the beginning

I’ve always loved mathematics. When I was a kid, my entertainments were mathematical: trips to the library brought back books like George Gamow’s (1947) “1, 2, 3, ... infinity” and even on recreational outings I would find myself doing something like estimating the number of windows in a (large, New York City) housing development. Although my working-class parents desperately wanted me to become a doctor, I loaded up on math courses as an undergrad, and when I told the mathematics chair at Queens College⁴ that I wanted to change from my pre-med chemistry major to mathematics, he asked “What took you so long?”

There are many reasons to love mathematics. Part of what makes mathematics so special is that it’s not arbitrary; you can figure things out. I have no idea at what point my first conscious mathematical discovery was. Perhaps it was the observation that every time I added two odd

³ This is somewhat oversimplified, of course. These issues overlap, substantively and chronologically – e.g., my first assessment project began in 1991.

⁴ It’s worth noting that aspects of the social compact were in place when I was an undergraduate. Education was considered a public good. Tuition and fees at Queens College (Part of the City University of New York) for New York City residents were \$32 per semester. All the way into the 1990s, tuition and fees at the University of California were under \$2000. Then, in a massive shift, politicians came to consider higher education to be a *private* good – people with college degrees earn more money over their lifetimes – and tuition began to skyrocket. The result is the massive student debt that current graduates suffer from – a distinctly U.S. phenomenon.

numbers, the result *had to* be even. Perhaps it was something else. Whatever it was, it was magical. And, it was mine – I’d figured it out, I *owned* it! This wasn’t somebody’s rule, which I had to memorize; this was something I’d figured out, and I understood why it was true. In other fields I had to memorize things. Where did Ohm’s law come from, for example? Or biological taxonomies, or multiple and conflicting interpretations of Hamlet? In math, things made sense.

To me, that meant that math was fundamentally democratic. It was open for discovery. And, assuming you showed (proved) something correctly, it was true – period. Nobody could argue it away; no authority could declare otherwise. What fun, what power! I had the sense, long ago (see, e.g., Schoenfeld 1994) that much of mathematics could be learned via sensemaking and problem solving. It seemed to me that most curricular content could emerge as the result of well-structured investigations rather than being imposed from on high. And, I had no idea that there could be anything political about mathematics. If anyone could do it and own it, how could it be political? (I remember a chat with some Spanish colleagues back in the late 1980s, when one of them claimed that mathematics was inherently political. I was incredulous – in fact, his statement didn’t really “compute.”)

Graduate school and early professional life – the late 60’s and early 70s

Some time before I earned my Ph.D. in mathematics, I ran across George Pólya’s (1945) *How to Solve It*. I read the book with fascination. Page after page, Pólya described methods of problem solving. As I read through the book, my smile got wider. If I was doing the things Pólya said mathematicians do, then I must be a real mathematician! But then I wondered, why hadn’t I been introduced to these methods? Was mathematics a secret guild, where the price of entry was figuring such things out for yourself? (In a sense, the answer is yes; but it’s more complex than that.)

In any case, doing math was fun. So was teaching. After earning my Ph.D., I taught for two years as a lecturer at U.C. Davis. That was my first introduction to academic politics: I was advised by my senior colleagues that I was spending far too much time with my students and that if I wanted to have a successful academic career I should limit my office hours and either close the door or go home to prove theorems. I very much enjoyed my teaching and earned high teaching evaluations; but I was told that that could be seen as a kiss of death among my department colleagues. The choice framed for me was, am I a researcher or a teacher?

At the time, I’d been reflecting (as a total amateur) on my teaching and had written about useful classroom techniques (Schoenfeld 1977). I spoke with a biologist friend who was involved in educational efforts at Berkeley. She suggested that I chat with Fred Reif, a physicist who chaired an interdisciplinary group called SESAME (Search for Excellence in Science and Mathematics Education) at Berkeley. Fred convinced me that there was a future to cognitive science and education, so I took a postdoc at Berkeley. Basically, I did so on the basis of an informal expected value computation. On the one hand, I loved mathematics and I wasn’t bad at it. But the odds that I’d do something transformative in mathematics were very slim – the pioneers of the previous few centuries were hard acts to follow, and the field itself had existed for two thousand years. By contrast, mathematics education was in its infancy. *ESM* dates to the 1960s, *JRME* to 1970. When I did my postdoc at Berkeley from 1975-78, the field of cognitive science didn’t really exist. (The first issue of the journal *Cognitive Science* appeared in 1977.) So, there were opportunities to participate in the growth of the field from the very beginning, bringing together my love of mathematics and my wish to go deeper into understanding mathematical thinking and

teaching. In addition, I'd always felt somewhat guilty being a professional mathematician. Being paid for producing theorems felt like being paid for doing crossword puzzles. It was fun, but to what benefit? If research on problem solving made it more accessible, then there was a potential payoff for students in terms of teaching and learning. I was, of course, totally naïve about what it takes to have an impact on school systems. But, the opportunity to shape the emergence of a new field, to combine my love of mathematics with my love of exploring thinking and learning, and, if I was lucky, to have some influence on practice, was irresistible.

The early years: Problem Solving Research and Development, 1975-1990

For a number of reasons I began my research on mathematical problem solving at the college level. I thought about working with doctoral students (but did they really need my help?) or on niche areas like the Putnam exam (but to tell the truth, I wasn't great at that kind of problem). I thought that Pólya's problem solving strategies were pretty sophisticated, so college students (rather than secondary students) were probably the right audience; I started with upper division Berkeley students, just to be safe.

What I found very quickly was that my students – among the best and the brightest – were woefully unfamiliar with even the most basic problem solving strategies. They were smart, they were creative... and they had gotten as far as they had because they were very good at mastering the mathematics they were instructed to master. There was, not only for these students, but in general, an unspoken didactical contract: their teacher will establish the context and show the students what they are responsible for. Homework assignments will stretch the students a bit, but they are largely repetitive. Tests will, with the possible exception of problems designed to “reveal the A students,” reward students who have done their homework. Although they were referring to K-12 curricula, Lappan and Phillips (2009) tapped into at least a K-14 universal when they referred to the dominant mode of instruction as “demonstrate and practice.”

The net result was that students had little or no experience with problem solving, or what Mason, Burton and Stacey (1982) called “thinking mathematically.” I came to realize that my students were fundamentally deprived, in mathematical terms. I moved my problem solving courses down to the lower division level, so that my students – whether intending math majors or not – could at least experience a dose of mathematical thinking. If they planned to go on in mathematics, they should at least have a sense of what mathematical sensemaking looked like. And if they didn't, then there was at least as much reason to give them a sense that mathematics could be interesting and exciting. There's enough mathophobia in the world as it is.

The core aspects of my work on mathematical problem solving are well known (see, e.g., Schoenfeld 1985, 1992), so I'll summarize them briefly. Then I'll discuss what I found along the way.

The central work on problem solving evolved over a decade or so. I first focused on problem solving strategies, or “heuristics,” as identified by Pólya (1945, 1954, 1962, 1965). The key insight was that Pólya was right about the strategies – mathematicians do use them, having picked them up, idiosyncratically, from their experience. (Rough paraphrase: a technique used twice becomes a strategy.) But, the grain size of Pólya's descriptions was wrong: a strategy such as “exploit an easier related problem” was unteachable on its own terms because actually consisted of at least a dozen different sub-strategies for identifying easier related problems and exploiting their solutions. My research showed that the sub-strategies could be learned and that when students learned enough of the sub-strategies, they could implement “the strategy.” (Rough

analogy: if you learn to cook a range of vegetables, and starches, and a variety of meats, then you can put together a complete balanced meal.)

Interestingly, solving the sub-strategy problem created a new problem. Pólya had identified perhaps two dozen major heuristic strategies, a manageable number. But if learning each strategy entailed learning a dozen sub-strategies, then the challenge jumped by an order of magnitude. The difficulty isn't simply a *learning* challenge, although mastering hundreds of techniques rather than dozens certainly ups the ante; it's a *management* challenge. How in the world do you decide which technique to use, when you have hundreds at your disposal? (Rough analogy: if I give you a key ring with a dozen keys, the odds are you'll be able to open a door within a reasonable amount of time. You can try them all if need be. But if I give you a key ring with hundreds of keys, the odds of your success diminish substantially.) That led to the issue of metacognition, more specifically the issue of monitoring and self-regulation. The bottom line is that self-monitoring can also be learned. With appropriate attention to reflecting on progress during problem solving, students can get good at it.

I was interested in what helped students succeed and what caused them to fail. There was no good reason to ask students to solve problems for which they didn't have the relevant knowledge, so I chose problems for which the students had the appropriate backgrounds. At the time, plane geometry was a required high school course, so I could be confident that the students knew the basics. I gave my first-year college students a simple geometry construction problem – which, despite their knowledge, they all approached empirically. I pursued the issue for some years, ultimately having the students prove results that solved the construction problem just before I gave them the construction problem to work on. Amazingly, the students ignored what they had just proved and made conjectures that contradicted it. Those findings led to the study of mathematical beliefs and their origins. And that pursuit led me into the schools, where I observed both what was taught and why. My experiences in schools led me to consider a series of structural issues, starting with the role of curricula and assessment in shaping students' learning.

Given that I had uncovered the challenge of unproductive beliefs in geometry, I started by sitting in on high school geometry classes. In the 1980s New York State was one of three major states (the other two being California and Texas) that had state-wide testing, along with state-supported curricula designed or selected in concert with the tests. What soon became apparent were the ways in which testing deformed instructional practice. The New York State Regents exams had a very specific format, with 10 points awarded for “solving” (i.e., reproducing the proof of) two proof problems out of a dozen or so “required” proofs. What happened of course was that students memorized all the proofs, for a guaranteed 20 points out of 100 on the exam. The test also had one “construction problem.” Students could earn two points for producing a sequence of lines and arcs on the page that looked just like one of the “required” constructions.

The way that instruction was organized in the school that I observed made the power of the exam very clear. Although geometric constructions were discussed about half-way through the text, the math department reorganized instruction so that constructions were taught just before the statewide exam. The rationale was simple: since students were intended to memorize the constructions and carry them out precisely, it was unwise to have too much time pass between memorizing and test-taking. Indeed, one of the most memorable quotes from instruction that year came from the teacher, shortly before students were to take a unit test on constructions: “you'll have to know all your constructions cold so that you don't spend a lot of time thinking about them.” The emphasis was on speed and accuracy, tailored to test performance. What mattered

when producing constructions was that the arcs on the page looked good, and that they were reasonably accurate.

A range of research findings included those observations (see, e.g., Schoenfeld 1988, 1989). These findings were not about any particular teacher; they were general. Hugh Burkhardt's acronym WYTIWYG ("what you test is what you get") accurately summarized the influence of high stakes assessments. That hasn't changed. Course texts were, and still are, tied to assessments. The literature has long indicated that teachers follow texts with great fidelity.

It should be stressed that the teachers in this and other studies were uniformly well intentioned – they were doing what they thought was in the best interests of their students. But, test pressures are enormous. That was the case even before the No Child Left Behind act, and it remains so. I have had National Board Certified teachers tell me that they would try out the ideas in our professional development program for one year, but if their students' test scores dropped by even one point, they would leave. I have seen an equity-focused teacher who built a summer program based on ideas related to growth mindset that was designed to help prepare "low-performing" students build confidence and agency completely forsake those ideas during the regular school year because there wasn't time for such things in a curriculum aimed at the high stakes state exams.

The point here is that by the early 1990s mathematics education researchers had a good idea of what mattered in mathematical performance. Understanding content – having mathematical resources at one's disposal – had always been considered important. NCTM endorsed "problem solving" and we had a theoretical understanding of how to decompose and teach heuristic strategies, although the process had not been done and curricula supporting problem solving instruction had not been built. The roles of metacognition and belief systems were understood, as were the causes of counterproductive beliefs. (Schoenfeld 1985, 1992). The obstacles to bringing these ideas into the classroom were structural and (socio)political.

The 1990s, and the math wars.

If there is one phrase to describe the 1990s in mathematics education, it's "the math wars." I've written extensively about this (Schoenfeld 2004, 2008; Schoenfeld & Pearson 2009) so I won't repeat the details but will make some observations.

People have multiple reasons for aligning with or leading political "movements," as has become all too clear in the intertwining of White supremacy, structural racism, fervid and sometimes "post factual," (e.g., QAnon) belief, and personal advantage in Trumpian politics. The same was the case, albeit not as blatantly, in the math wars. There is no question that some of the participants considered the integrity of mathematics to be at stake and felt that they were protecting it. There is also no question that partisanship gave some people, both inside and outside the mathematics community, opportunities for personal advantage such as visible prominence and political advancement. Here I want to point to some more structural issues.

The first issue is financial. I was a member of the group that wrote the 1992 California mathematics framework. Our meetings were public. They were sparsely attended, with one exception. There were always representatives from major publishers at our meetings. They delivered one clear message: "reform" is impossible because it would be too expensive. It cost \$25 million to develop and produce a K-8 textbook series, they said, and no publisher was going to risk that much money on an untried concept. They were right. What did happen was that the

National Science Foundation realized that the lack of suitable textbooks was a roadblock to progress, so NSF issued a funding initiative for the production of “standards-based” or “reform” texts.

Reform texts catalyzed the math wars, which raged over much of the 1990s. To understand whether politics or substance matters, it is essential to note that the math wars were waged largely in the absence of hard data. The motivation for reform was clear: there was undeniable evidence of the shortcomings of “traditional” instruction and a decade of small-scale reform-oriented studies suggested that the directions in which the NSF-funded curricula were headed were likely to be productive. The hard evidence to support this hypothesis didn’t really start coming in until 2000, however. The case for reform became stronger when Senk and Thompson’s (2002) summary volume indicated across-the-boards wins for reform. (The one-line summary: students using standards-based materials did roughly the same on tests of skills as students who received “traditional” instruction; they outperformed such students on tests of problem solving and conceptual understanding. See Schoenfeld 2002). The fact that the wars persisted for so long in the absence of hard data, indicates that the forces that drove much educational policy were political rather than grounded in data and research.

Moreover, although the issues are far less direct, issues of race and its oftentimes inextricable partner, socioeconomic status, were also implicated. One of the underpinnings of the Standards movement, and an explicit goal of some Standards-based curricula, was to move toward more equitable instruction⁵. To state things directly, there are those who believe that excellence and equity are in conflict – that there is a gradation of mathematical talent, and that an attempt to enfranchise all students mathematically is a disservice to those talented students who would profit from more “rigorous” training. From that perspective, if less “talented” students fall off the mathematical ladder, that’s their problem; there is “enough” mathematical talent to advance the nation’s interests, and one should not dilute instruction to serve the masses.

The math wars were fomented in California by a group called “Mathematically Correct,” whose website still exists at <http://www.mathematicallycorrect.com/>. It is no accident that the wellsprings of Mathematically Correct were San Diego and Palo Alto. San Diego was a hotbed of right-wing conservatism, partly because of its proximity to the Mexican border and the fact that the Spanish-speaking population was increasing rapidly. Immigration backlash included the sponsoring of California’s state proposition 227, essentially an “English only” mandate for the schools. The analysis from “Ballotpedia,” an independent analysis group, summarized Prop 227 as follows⁶:

“According to Ballotpedia, “Proposition 227 changed the way that “Limited English Proficient” (LEP) students are taught in California. Specifically, it

- Required California public schools to teach LEP students in special classes that are taught nearly all in English. This provision had the effect of eliminating “bilingual” classes in most cases.

⁵ NCTM’s stance on equity and diversity has been problematized. See, e.g., Martin (2009); for an update see Martin (2019). However partial or inadequate the NCTM position may have been, it was a flash point for controversy, as discussed above.

⁶ Details taken from Wikipedia, https://en.wikipedia.org/wiki/1998_California_Proposition_227.

- Shortened the time that most LEP students stayed in special classes.
- Eliminated most programs in the state that provided multi-year special classes to LEP students by requiring that (1) LEP students move from special classes to regular classes when they had acquired a good working knowledge of English and (2) these special classes not normally last longer than one year.

This was one of many policies implemented by the right wing, tapping into the xenophobia that ultimately metastasized into Donald Trump’s immigration policies. In San Diego in the 1990s the White voting majority felt threatened by a growing Latino minority. Claiming that high quality learning was being threatened by untested equity-driven mathematics programs, with the tacit implication that the new programs were tailored to minority students, was a perfect wedge issue to mobilize White voters.

Palo Alto represented similar issues in a different way. The area had a mixed but separated demographic: Palo Alto itself was upper-upper middle class, and East Palo Alto had a largely minoritized population consisting in large measure of the people who cooked for, cleaned for, and maintained the homes of those (literally!) on the other side of the tracks. Here the issue wasn’t fear of disenfranchisement, but loss of privilege. The “good schools” in Palo Alto reliably sent their students to the best schools and universities. Why tinker with success, for abstract reasons of equity and diversity? If it ain’t broke (for your children, that is), don’t fix it. (N.B. My rhetoric is mild, but the rhetoric of the political battles in Palo Alto was anything but.)

The tensions remain. The same right-wing players who brought us the math wars are now manufacturing a battle over the anti-tracking stance in the 2021 draft California Mathematics Framework.

The 2000s and No Child Left Behind

On a purely personal note, I want to bring up one of my signal failures. I led the secondary standards writing team for NCTM’s (2000) *Principles and Standards for School Mathematics*, the successor volume to the 1989 NCTM *Standards*. The process by which *Principles and Standards* was created and vetted was beautifully managed and the endorsement of the process by all of the major mathematics societies quieted the math wars.

All too aware of the impact of testing and the import of WYTIWG, I argued that our assigned task, writing standards and providing examples of interesting classroom activities, was good but not good enough: the wording of the Standards (on the order of “students will understand X”) was somewhat vague and aspirational, and could be misconstrued. At the time the standards movement nationwide was morphing into a test-based accountability movement, so the nature of assessments was vitally important. It would be easy, I argued, to craft assessments that bastardized our intentions. I proposed to the writing team that we incorporate sample assessments into our document. It was put to a vote, and I lost 25 to 1.

The result was a disaster. Had NCTM produced sample assessments, it would have taken the lead in saying “This is what we want students to be able to do with the mathematics they learn;” there is a good chance that such tests would have shaped statewide assessments, and thus shaped instruction. When NCTM failed to do so, the vacuum was filled by the No Child Left Behind act (commonly known as NCLB; see US Department of Education 2002). Under NCLB, states built their own assessments. Most of those assessments, in line with traditional assessments, focused on low-level skills. This effectively undermined the goals of *Principles and Standards*.

Some good intentions motivated the creation of NCLB. Each state defined its own standards, assessments, and performance targets. The idea was to ratchet up performance standards gradually and to provide support and rewards for reaching those standards. Disaggregation mechanisms – *every* demographic group had to meet the standards – assured attention to the performance of all students. And, there were carrots and sticks. The carrots were that schools that lagged behind would be given significant resources to improve. The sticks were that if they failed to improve for “too long,” penalties would be imposed. Individual students would be left back; teachers would be dismissed; if a school failed to meet progress goals for a number of years in a row it would be dismantled, and whole districts would be put in receivership.

That approach might have been workable (although highly punitive) if the carrots were in place, so that districts that faced challenges in making adequate progress were provided with resources to address the challenges they faced. But guess what? In the congressional sausage-making process, the penalties for failing to make progress were carved in stone but the resources to support failing districts were never authorized.

Problems abounded. There was huge variability in the sets of standards and assessments built by the states. Most of the assessments were of low quality. Some states gamed the system, demanding minimal progress until right before the 2013-14 year; somewhat paradoxically, setting low standards allowed them to avoid penalties. But the main issue was structural inequality. Rich districts had the resources to do well enough, at least at the beginning. That’s not to say that testing didn’t bend schools out of shape. When I went to observe some classes in late February, a teacher I knew from my having worked in her school told me not to bother watching instruction – “all we’re doing is prepping for the test, there’s no real teaching going on.” But that was a district that could afford “business as usual.” The challenges faced by poor and minoritized districts were far worse.

Because the enacted version of NCLB failed to provide fiscal support for “failing” districts but did penalize them, under-resourced and minoritized districts quickly found themselves being penalized. A local district I was working in was the first to be put in receivership. The results were devastating, adding injury to injury. The district, with a 90% minoritized population and hardly any resources, was forced into continuous test-prep mode. The result is that the students who were in the greatest need of meaningful instruction were systematically denied it. This was but one example of the structural racism within the system. It is consistent, of course, with many equally blatant examples (e.g., Kozol 1992, Rothstein 2107).

The bottom line: seemingly reasonable policy decisions can have significant negative impact on people’s lives. This, again, is the issue of “learning loss.”

The 2000s, Part 2: The What Works Clearinghouse

Some background on testing is necessary before I proceed here. Testing is *not* a neutral measure of proficiency. Any test assesses what is declared to be important to some degree, depending on how artfully the test is constructed. But there’s great variation. Take literacy as an example. If you define “literacy” as having a specific vocabulary, you give vocabulary tests and teachers wind up drilling their students on vocabulary. If you define “literacy” as the ability to analyze text, you develop a very different kind of test; kids read and think. The nature of the test is consequential, because students are declared to be “literate” (or not) based on their test scores.

It's the same in math. One definition of mathematical proficiency is skills- and knowledge-based. The basic idea is that students should be able to execute the skills taught in the curriculum. A more inclusive definition calls for having students demonstrate proficiency in skills, conceptual understanding, and problem solving. Depending on which approach you assess, you get very different results.

Table 1, taken from Ridgway, Crust, Burkhardt, Wilcox, Fisher, and Foster (2000), shows the differences in the two approaches. The SAT-9 was a skills-based assessment used across California in the 1990s. The MARS test was a test of skills, concepts, and problem solving. Ridgway and colleagues (2000) administered both tests to more than 5000 students at each of grades 3, 5, and 7. The patterns are the same across grades.

MARS	SAT-9	
	Not Proficient	Proficient
Grade 3 (N = 6136)		
Not proficient	27%	21%
Proficient	6%	46%
Grade 5 (N = 5247)		
Not proficient	28%	18%
Proficient	5%	49%
Grade 7 (N = 5037)		
Not proficient	32%	28%
Proficient	2%	38%

Let's take grade 3 as an example. If a student was declared proficient on the MARS test, there's a $46/52 = 88\%$ chance that the student would be declared proficient on the SAT-9. That looks like pretty good alignment. But if a student was declared proficient on the SAT-9, there was a $46/67 = 69\%$ chance that the student will be declared proficient on the MARS test. To put this more directly, 31% of the students declared to be "proficient" by California's official test turn out to be "not proficient" when conceptual understanding and problem solving were assessed in addition to skills.

This is a big deal if you care about conceptual understanding and problem solving. It's also a big deal if you care about curricula. If you use the SAT-9 to assess grade 3 performance, it looks like 67% of the students are proficient. That's not wonderful, but it seems within bounds; there appears to be a base for improvement. If you use the MARS test to assess grade 3 performance, however, you get a very different story. Only 52% of the students test proficient; you'd better make some radical changes. In short, *what you test matters*. It's shocking that I have to say this – but read on.

The background just provided establishes the context for some general comments and then a description of my specific experience with the "proficiency-based" testing.

No Child Left Behind was only one of the educational policy initiatives put in place during George W. Bush's presidency. There was also the misguided attempt on the part of the U.S. Department of Education to define randomized controlled trials (RCTs) as the "gold standard" of

educational research (see, e.g., US. Department of Education, 2003). The issue is not that randomized trials aren't an excellent way of conducting research under certain circumstances. The issue is that in pragmatic terms, the Department of Education discounted almost all other forms of research – it didn't consider evidence produced by alternative methods to be adequate evidence of effectiveness. For a broader discussion of alternative methods and their validity, see *Scientific Research in Education* (National Research Council, 2002), which was produced in rebuttal to the Department of Education's agenda (although, of course, it didn't say so); see also Schoenfeld 2007, which lays out criteria for rigorous and meaningful research and problematizes the use of randomized controlled trials in educational research.

If this discussion were merely “academic,” that would be one thing. But the relevant issues turn out to be Political, with a capital P. As part of its agenda to certify high quality instructional materials, the Department of Education's Institute of Educational Sciences created the “What Works Clearinghouse.” WWC's mandate was to certify when instructional interventions had been validated by rigorous means. If a curriculum or other instructional treatment had been evaluated by means of some formal assessment, WWC staff would evaluate the quality of the evaluation. A carefully conducted randomized controlled trial would get you the highest marks – the equivalent of the “good housekeeping seal of approval” for the curriculum. WWC was established to conduct a large literature review, identifying interventions that “worked.”

In 2003 I was appointed the WWC's Senior Content Advisor for Mathematics. Basically, my responsibility was to ensure the intellectual integrity of the enterprise. (Staff did the work.)

An early article I produced for WWC was intended to serve as part of a large technical document that framed WWC's approach to certifying instructional materials. The article explained the history of mathematics curriculum development and assessment. It predicted slim pickings for the WWC mathematics literature review, because very few instructional treatments had been subjected to the kinds of randomized controlled trials that the WWC used as its evaluation standard. The Institute for Educational Sciences (IES), which funds WWC, instructed WWC to remove my article from the document. When I complained, WWC said that I would have the opportunity to revise the document for publication when some instructional treatments had been reviewed and WWC was further along in the process.

I waited. After the WWC staff produced its first series of evaluations, I was informed that the Clearinghouse planned to work with a journal to create a special journal issue characterizing WWC's work. I was told to update my article and submit it to the journal. When I went through the new data, the predictions I had made in my earlier piece were confirmed: very few studies met WWC's criteria.

More importantly, as I worked through the data I discovered a fundamental flaw in WWC methodology. WWC had not analyzed what the assessment measures used in the studies actually assessed. Thus WWC's certifications of quality had little meaning. When an instructional treatment was judged to meet WWC criteria, it was impossible to know what exactly the treatment did well. Did students learn skills, or problem solving, or conceptual understanding, or something else? There was no way of knowing without conducting a content analysis of the assessment. Because there was only a handful of certified programs, I urged WWC to conduct the relevant content analyses. They refused.

I revised my article. The revision, like its antecedent, contained the history of mathematics curricula and my prediction of slim pickings. It documented the accuracy of the predictions and

contained a discussion of why WWC’s refusal to perform content analyses was deeply problematic. I submitted my revision to the journal and waited for reviews. After a very long delay I was informed that, after conducting a “prepublication review,” IES had instructed WWC to remove from the journal every single paper that had been written by WWC staff. Of course, it made no sense to publish my paper as a stand-alone; the journal issue was cancelled.

The only way I can see to interpret the sequence of events that I have just described is that IES killed the special issue in order to prevent the publication of my piece. (This is not the first case of a federal agency blocking publication of “inconvenient truths.” There is a history of such actions with regard to climate change and other areas.) I resigned from my role as a WWC advisor and published the details of the story in *Education Researcher*. For that detail, See my *ER* article (Schoenfeld 2006a), WWC’s rejoinder (Herman et al. 2006) and my response (Schoenfeld 2006b).

The 2010s

The previous section described the negative impact of deliberate policy choices. What follows in this section features the laws of unintended consequences – the epitaph for which is a Robert Burns’ quote, “The best laid schemes o’ mice an’ men / Gang aft a-gley.” I include these stories because they help frame my final discussion of the goals of mathematics instruction.

In math-ed terms, the 2010s can be considered the decade of the Common Core and the assessment systems that enforce it. Despite the best of intentions and some positive outcomes (e.g., greater consistency in nation-wide goals for instruction) there are deeply problematic aspects to both. I appreciate the challenges faced by the authors of the Common Core, who had little time to compile their work and were doing their best to avoid re-kindling the math wars. The result, in contrast to the NCTM *Standards* volumes, is a rather slender volume. NCTM’s (2000) *Principles and Standards* weighed in at more than 400 densely packed pages that described and exemplified content and processes, with equal space given to both – that is, the fundamental *processes* of problem solving, reasoning, communicating, making connections, and using mathematical representations received as much attention as the *content* that was described (number and operations, algebra, geometry, measurement, and data analysis and probability).

The Common Core contains a 3-page list of “standards for mathematical practices” and a 74-page list of “standards for mathematical content.” Here is a sample from the beginning of the Grade 6 content description.

[CCSS.Math.Content.6.RP.A.3](#)

Use ratio and rate reasoning to solve real-world and mathematical problems, e.g., by reasoning about tables of equivalent ratios, tape diagrams, double number line diagrams, or equations.

[CCSS.Math.Content.6.RP.A.3.a](#)

Make tables of equivalent ratios relating quantities with whole-number measurements, find missing values in the tables, and plot the pairs of values on the coordinate plane. Use tables to compare ratios.

[CCSS.Math.Content.6.RP.A.3.b](#)

Solve unit rate problems including those involving unit pricing and constant speed. *For example, if it took 7 hours to mow 4 lawns, then at that rate, how many lawns could be mowed in 35 hours? At what rate were lawns being mowed?*

[CCSS.Math.Content.6.RP.A.3.c](#)

Find a percent of a quantity as a rate per 100 (e.g., 30% of a quantity means 30/100 times the quantity); solve problems involving finding the whole, given a part and the percent.

[CCSS.Math.Content.6.RP.A.3.d](#)

Use ratio reasoning to convert measurement units; manipulate and transform units appropriately when multiplying or dividing quantities.

In my experience, such neat clean descriptions get turned by school districts into curricular checklists – “Have we worked problems on tables of equivalent ratios? Yes, check. Have we practiced on tape diagrams? Yes, check. And so on. That is, it’s easy to go from bullet points to scope and sequence. You get content “coverage” in the narrowest sense.

And what about important practices and processes? The Common Core discusses 8 key practices on pages 6-8, and the list of those practices is reprinted at the beginning of each content chapter. We’ve known for decades that problem solving success hinges on: students’ knowledge base; their access to problem solving strategies; effective metacognition, specifically monitoring and self-regulation; and productive belief systems, about mathematics and about oneself vis-à-vis mathematics – in today’s language, productive mathematical identities. (Schoenfeld 1985). To speak bluntly, the Common Core offers no meaningful support for anything but content. Functionally, there is no support for problem solving. There is no support for the development of metacognitive skills, and beliefs and mathematical identities are not addressed.

When discussing “real world” implementations of curricula, Ann Brown observed that all curricula and frameworks undergo mutations when they move into classrooms. The challenge, she said, was to avoid lethal mutations. Unfortunately, (a) the Common Core’s list of bullet points is easily converted into a curricular scope-and-sequence. (b) The Common Core offered no meaningful support for mathematical processes and practices and (c) it provided no exemplification of rich and interesting mathematical problems and discussions. In consequence, in addition to not providing direct support for ambitious curricula, the Common Core left the door wide open for lethal mutations.

There are two ways out of this dilemma. The first is large-scale curriculum development, a process that takes many years and large investments. For the most part, that just didn’t happen. It is the case that some good Standards-based curricula were retrofitted to the Common Core, and some ongoing projects are providing good materials. The problem is that high stakes assessments were going to be implemented soon after the Common Core was adopted. School districts needed Common-Core-consistent curricula ASAP. The results were mostly cut-and-paste disasters. This is a systemic failure.

The second way out of the lethal mutations dilemma could be the use of well-constructed assessments. Given WYTIWYG (What You Test is What You Get), a set of robust assessments that interpreted the Common Core in the right ways could have driven instruction in the right directions.

Hugh Burkhardt and I were asked to head the team that drafted the specifications for the Smarter Balanced Assessment Consortium (Schoenfeld & Burkhardt 2012), which contracted with about half the states in the US to implement Common-Core-consistent assessments. We were excited about the possibilities because they promised two fundamental changes. First, we constructed a system that was able to provide meaningful and reliable sub-scores regarding students’

knowledge of (1) concepts and procedures, (2) problem solving, (3) communicating reasoning, and (4) modeling and data analysis. The point of such sub-scores is that they can highlight particular students' or schools' strengths and weaknesses rather than providing single numerical grades. Second, we exemplified the assessment with a collection of mathematical tasks that embodied the mathematical richness we wanted students to engage with. The tasks were taken largely from task banks constructed by the Mathematics Assessment Project (see <https://www.map.mathshell.org/>). They have been used in secure testing situations for many years.

The 2012 SBAC specs are no longer on the SBAC web site. In fact, the specs were never implemented in ways consistent with Burkhardt's and my expectations. The problem is that the SBAC governing Board always planned to move toward computer-graded exams, which are cheaper, more "reliable," and more "secure" than hand-graded exams⁷. SBAC built what it could and implemented it. In my opinion the transition to computer-based exams de-natured the mathematics in the exams to the point where the exams fail to represent the mathematical richness that we had built into the exam specifications. (To be fair to SBAC, I am measuring them against high standards. The assessments produced by the other national assessment consortium, PARCC, are far worse.)

The point is that there have been opportunities to orient the system toward richer and more engaging – even if somewhat traditional – mathematical content. For largely systemic and political reasons, that hasn't been done.

The 2020s

The first two years of the 2020s have already given us more than a decade's worth of challenges. As I wrote in Schoenfeld (in press),

“The murders of George Floyd, Breonna Taylor, Trayvon Martin, Sandra Bland, Ahmaud Arbery, and numerous other Blacks at the hands of police and white supremacists laid bare for all except those who refuse to acknowledge it the structural racism that underpins American society (Center for American Progress 2019, Urban Institute 2020, Wilkerson 2020). It's not that such issues were unknown; it's that the murders and the protests they engendered made them much more difficult to ignore. The reality that many minoritized people live in a world apart from White America, with different and much more devastating expectations for quality of life (including education) has been rendered day after day in high resolution.”

If anything, the situation has gotten worse in the year since I penned those words. The completely manufactured “controversy” over teaching Critical Race Theory in schools represents a full-fledged attempt to ban the teaching of the history of oppression described in the previous paragraph. Such actions and their consequences reach into every mathematics classroom.

In much of my previous work I theorized about what took place inside ostensibly closed systems – people solving problems in isolation, teachers making decisions, actions in the classroom. My problem solving research asked, “what are the aspects of thinking and understanding that need to be examined in order to determine the success or failure of any individual's attempt to solve a problem?” The (theoretically complete) answer was, “You need to know about the individual's

⁷ There are mechanisms for hand-grading exams that are comparatively inexpensive and secure – see the arguments in Burkhardt & Schoenfeld 2019. AI-graded exams could have been phased in gradually.

knowledge base, problem solving strategies, metacognitive behavior, and belief systems” (Schoenfeld 1985, 1992). Similarly, my research on teachers’ decision-making asked, “what do you need to know in order to model the in-the-moment choices a teacher makes during instruction?” The (theoretically complete) answer was, “If you know the teacher’s resources, orientations, and goals in very fine detail, then you can produce a detailed model of the teacher’s choices by using a specific probability-based decision mechanism” (Schoenfeld 2010). My ongoing classroom research asks the question, “which dimensions of classroom interactions are necessary and sufficient to ensure that students will emerge from instruction as knowledgeable, resourceful, and agentic thinkers and problem solvers? The (theoretically complete) answer is, “it suffices to examine the five dimensions of the Teaching for Robust Understanding (TRU) Framework: the quality of the mathematics; opportunities for productive struggle; equitable access to meaningful engagement with core content; opportunities for the development of agency and positive mathematical identities; and formative assessment” (Schoenfeld 2013, 2104; Schoenfeld & the Teaching for Robust Understanding Project 2016).

The challenge is that these closed systems, while allowing for theoretically complete solutions, also wind up finessing key questions of causality. In problem solving, where do knowledge and belief systems come from? In a society where students are stereotyped, tracked, and provided very different opportunities to learn, such issues matter: it’s not just what individuals bring to a problem situation, it’s what shaped their knowledge and belief systems before they sat down to work on the problem. It’s the same with teaching: Where do teachers’ orientations come from? For example, we have to think about what led a teacher to say, when I asked him whether he’d ever consider giving his class a problem and let them grapple with it, “not these students, it would just confuse them. I do that with my honors students.” And, when we think about the construction of powerful learning environments along the lines of the TRU Framework, we have to think about the distributions of opportunity to do so. These are massive societal issues. The solutions to the closed system problems point to what needs to be done, but larger systemic issues need to be taken into account when we consider what caused things to be as they are and how we might address them.

Nowhere is the set of larger social issues clearer than when we consider the national impact of Covid-19. Essential workers – disproportionately people of color – were forced to work but were not given prioritization for vaccination. The consequences are all too predictable, as indicated by a piece in the *New England Journal of Medicine* entitled “Structural Racism, Social Risk Factors, and Covid-19 – A Dangerous Convergence for Black Americans” (Egede & Walker, 2020).

Similarly, children of poverty and children of color suffer the academic impacts of Covid disproportionately: see “Addressing Inequities in Education: Considerations for Black Children and Youth in the Era of COVID-19” (Gaylord-Harden, Adams-Bass, Bogan, Francis, Scott, Seaton, and Williams 2020). That article has the following section heads:

- Systemic Racism is the Pre-Existing Condition Affording COVID-19 the Opportunity to Disproportionately Impact the Black American Community;
- Black Families are Facing More Severe Economic Consequences;
- Black Children Face Disadvantages in Remote Learning Settings;
- Schools That Serve Black Children are Less Able to Provide Remote Learning Experiences; and
- Black Children are Experiencing Elevated Levels of Stress.

There are similar findings for a wide range of minoritized populations; see the collection of SRCED policy papers at <https://www.srcd.org/research/briefs-fact-sheets/statements-evidence>.

Despite consistent evidence along these lines, the predominant concerns in the media are focused on “learning losses.” That’s what makes headlines. Earlier this year I wrote an editorial to that effect, titled “It’s Time for an Academic Reset.” It made the following arguments:

What really matters? First and foremost, students’ mental and emotional well-being. COVID’s impact has fallen disproportionately on communities of color and people who are economically disadvantaged. Privileged students often have good technology, good wifi, and nice places to study. One of my former students, who teaches in a low-income, highly diverse district, had to find her students to give them electronic tablets they could work on; then some of those students had to park themselves outside of schools to get a wifi signal. The current crisis magnifies longstanding inequities. Making believe we can make “normal” progress under these circumstances without doing serious damage to the most disadvantaged students is just plain crazy. We need to find modes of schooling that support students socially and academically. (See <https://gse.berkeley.edu/news/its-time-academic-reset> for the editorial. See also McKinney de Royston & Vossoughi, 2021)

The editorial was rejected by the New York Times, the Washington Post, the Sacramento Bee, Education Week, and more. Well, OK, maybe they weren’t interested in the arguments put forth by a lone academic. So, I worked with the Laureate Chapter of the education honor society Kappa Delta Pi – a who’s who of scholars and equity advocates – to craft an updated version of the editorial. No luck. The challenge is that “Learning loss” sells in policy terms, while thoughtful examinations of underlying issues are a tough sell. And inequities persist.

What really matters in mathematics learning?

In this section I put forth a somewhat radical proposal based on (a) reflections on my recent experiences using mathematics in my personal life, (b) recent political events, and (c) reflections on the evolution of ways in which I have been teaching my course on mathematical thinking and problem solving.

But, let me start with my roots. I’m a math person. My Ph.D. is in topology and measure theory and I truly love mathematics. I have spent my entire professional career aimed at the goals described in the opening paragraph of this chapter: “Can we understand enough about mathematical thinking and problem solving to help students get good at it? Wouldn’t it be great if increasing numbers of students could experience the power and beauty of mathematics, and even come to feel about it the way I do?”

Over the years, my thinking about what matters has broadened. There are problem solving strategies; there are issues of monitoring and self-regulation; there are belief systems. There are what I’ve called “productive patterns of mathematical thinking” (Schoenfeld 2017) or, more traditionally, mathematical practices (Schoenfeld 2020a). Then there are questions of what kinds of learning environments support students in developing such understandings, and what it takes to teach for robust understanding of mathematics (Schoenfeld 2020b; Schoenfeld and the Teaching for Robust Understanding Project 2018.) So, my roots are firmly planted in (relatively pure) mathematical soil. But...

When I ask the question “what mathematics have I used in my non-professional life that was important and consequential?” the answer is “almost nothing I learned in school.” And yet, I have made very meaningful use of straightforward mathematics. Here are two examples.

Example 1. I chair the Coronavirus Advisory Committee (CAC) for a residential program that serves adults who have developmental and other disabilities. CAC is responsible for setting policies and protocols for residents and staff that concern vaccinations, safety, masking and distancing, travel, and testing. Establishing and updating these policies takes place in the context of rapidly changing and often incomplete or contradictory information and recommendations from available sources. When you look closely, it becomes clear that some policy decisions, including those from US government agencies such as the FDA and CDC, are politically influenced. Indeed, within a period of weeks, new guidance issued by these agencies has conflicted with earlier recommendations without the evidence base having changed substantially. This is a “real world” problem of some significance. How do you think about issues of Covid rationally, based on available information? How do you cut through conflicting information to make sane policy decisions? Here’s a problem, which I discuss in Schoenfeld 2021:

It is now generally accepted that the primary mechanism of Covid-19 transmission is the inhalation of aerosol particles; Under most circumstances 6 feet of physical distancing is considered a safe distance to avoid infection. Let’s take those as scientifically established for the sake of discussion. The other day as I was out for a walk (wearing a cloth mask) I was irritated by cigarette smoke produced by a smoker who was across the street, a good 30 feet away. If an aerosol irritant could bother me at a distance of 30 feet, why is 6 feet of physical distancing considered safe for COVID? (Schoenfeld 2021, p. 397)

Have fun with this problem if you wish (or, see my solution in Schoenfeld 2021). In broad-brush terms, here’s how I thought about the problem. I don’t know much biology, but that’s not an issue regarding this problem: if I could frame the underlying issues in the right ways a google search would give me reasonable data. What I needed to do was figure out the right questions to ask. These questions were all I needed to address the issue. Regarding Covid transmission: how big are infectious Covid particles and how far are they likely to travel? How dense are they in an infected person’s exhalations? Similarly regarding cigarettes: how big are cigarette smoke particles and how far are they likely to travel? How dense are they in a smoker’s exhalations? Answers to those questions were easy to find and to triangulate. Once I had them, some elementary mathematics resolved the issue. (Smoke particles waft, and there are tons of them. There are way less covid particles, which are much larger, and sink.) This type of thinking with emerging data has helped our Coronavirus Advisory Committee establish and modify appropriate safety protocols.

Example 2. This case of mathematical thinking concerns my personal health. I was diagnosed as having type 2 (adult onset) diabetes more than 20 years ago. It’s not a major concern; my blood sugar levels are easily kept within bounds by a combination of diet, pills, and exercise. When I was first diagnosed I started keeping track of what I ate and how my blood sugar levels changed. I quickly learned that the general dietary guidance provided by nutritionists is of limited use because the dietary categories in the recommendations are too broad and there are significant differences in metabolism from individual to individual. White rice sends my sugar skyrocketing, for example, but brown rice is fine; my favorite Chinese restaurant noodle dish sent my sugar through the roof and I had to stop eating it, while my homemade pasta wasn’t a problem. Simple data tracking revealed which of my pleasures I could enjoy without significant risk. It also

revealed, contrary to dietitians' dogma, that a reasonable quantity of wine with dinner of wine lowered my average blood sugar rather than raising it. To settle a longstanding point of contention with my doctor, I went for 3 weeks without wine and compared my sugar levels with those of the previous 3 weeks. Wine won over abstinence!

A more serious issue arose recently when my doctor suggested substituting a new diabetes pill (medicine A) for a pill I'd been using (medicine B), because the newer medicine offers increased protection against heart disease. To my dismay but not my surprise, no information was available regarding how doses of medicine A and medicine B compare. So, my doctor and I had to proceed empirically.

Medicine A comes in doses of 10 and 25 mg. Our first empirical trial involved a roughly half-and-half switch: I added the small (10 mg) dose of medicine A to my daily regimen and cut back half on B. (To give my metabolism time to stabilize, each of the empirical trials described here took about 3 weeks.) The numbers from the half-and-half switch looked pretty good.

The next question was, is 10 mg of A enough by itself? To find out I stopped taking B. That didn't work well; my blood sugar rose above the levels we wanted. That led us to consider the 25 mg dose of medicine A. Under the natural assumption that the impact of A would be proportional to dosage, we expected the extra 15 mg of A to provide a very good reduction in blood sugar levels. So, I stayed off medicine B but increased medicine A to 25 mg. The result was a surprise. There wasn't nearly as much effect as we expected – the 25 mg of A didn't reduce my blood sugar levels much more than the 10 mg dose had.

That meant that we had to reconsider our basic model. My doctor had said that medicines A and B used two different mechanisms to remove sugar from people's systems. Since the move from 10 mg of A to 25 mg of A didn't help that much, it was now reasonable to assume that the mechanism by which A worked had maxed out at a little more than 10 mg. On the other hand, since medicine B worked by a different biological mechanism, its impact might well be *in addition to* that of medicine A. That's why 10 mg of A plus half of the B I'd been taking had been effective.

I won't run through the numbers here, but I will say they're compelling. What I want to focus on is the process that produced the results. My doctor and I faced a situation for which there was no medical guidance, but for which short-term experimentation was low risk (I could stop taking any combination of medicines immediately if my blood sugar numbers looked bad). We did some simple experiments assuming that the impact of the drugs would be proportional to the dosage, and then revised our assumptions when the data didn't turn out as expected. The result is a much better medical regime for me.

The kind of thinking described in examples 1 and 2 could literally be matters of life and death. In both cases I wondered if the situation at hand could be modeled using some simple proportional reasoning. And – and this is the critical part – in both cases I had the sense of agency that led me to build the models and see if they explained things.

The odds are that a very small percentage of people would think in these ways or have the personal agency to do this kind of mathematically based experimentation. That's a very big problem.

I believe that problem comes in large part from the insularity of the curriculum and from the lack of agency that students develop because of the ways we teach. By insularity, I mean that students

historically learn to solve only the categories of problems we explicitly prepare them to solve. Rather than thinking of the mathematics they've learned as tools that could apply in a wide range of situations, they think of that mathematics as applying to very narrow classes of problems – specifically, the kinds of problems they've been taught to solve. Just as students learned to expect that “all problems can be solved in five minutes or less” on the basis of their classroom experience (see Schoenfeld 1985), students also learn to expect that “the math I learn in school is not applicable in meaningful ways to issues that take place outside the classroom.” With such expectations, they don't think to use the math they know in situations like those in examples 1 and 2. The problem is compounded by the fact that most students have almost no experience pursuing mathematical ideas on their own. If you haven't done so in the classroom, why would you do so outside the classroom?

Mathematical agency is a fundamentally important issue. I am, once again, teaching my problem solving course this semester. Over the years I've found myself “covering” less and less, in that my students and I work fewer problems than before – but we work them much more deeply, exploring the mathematical issues and connections they might suggest. This semester my students and I were playing with the mathematics of 3×3 magic squares. In looking at possible extensions and generalizations a student conjectured that the sum of 9 consecutive integers would always be divisible by 3. That student ultimately argued that (a) the sum of 3 consecutive integers could be shown to be divisible by 3; (b) 9 consecutive integers could be divided into 3 triples, each of which is divisible by 3; (c) since 3 was a factor of each triple, 3 was thus a factor of the sum.

The student's observation our reflections on it led to other questions. What about the sum of 5 consecutive integers? What about the sum of n consecutive integers, if n is odd? What if n is even? Things got complicated as we played with examples. Some numbers could be obtained as sums of consecutive integers, but some (4 and 8, for example) couldn't. That led to this question: which integers *can* be expressed as the sum of consecutive integers? The class was off and running, in directions I hadn't expected. They worked through the class break, wrote about the problem passionately in our class logs, and ultimately followed their ideas until they produced a complete solution to the problem. Now, in this particular instance my students produced a solution to a known problem, but that doesn't matter⁸. They were doing mathematics, and it was exhilarating. More important than the fact that they solved a particular problem was the fact that they saw themselves as honest-to-goodness mathematical sense makers. When you have that sense of yourself, you're empowered to tackle new problems – and if doing so becomes enough of a habit, you might feel empowered enough to take on the kinds of Covid and health-related problems I discussed at the beginning of this section. That is: students who have engaged in that kind of generative mathematical thinking throughout their academic careers are much more likely to be mathematically agentic.

Let me try to pull the various themes of this discussion together. First, structural inequities in schooling have worsened during covid. Society at large has done its best to ignore the issue, focusing on meaningless “learning loss” instead. Second, we know that myriad students are disaffected from mathematics. There are multiple reasons for this, including its perceived irrelevance and perceived inaccessibility. Third, if people can't use elementary mathematics to

⁸ Historical note: My students have, at times, derived new mathematics when pursuing ideas they found interesting, and the results have been published. That didn't happen this semester, but that's not the point. What matters is that these students saw themselves as capable of creating new mathematics and took great pleasure in it.

reason about what are literally life-and-death issues, mathematics as taught is a dismal failure. Fourth, if people have no mathematical agency, they won't use what they "know," so their school knowledge is irrelevant.

If you take these issues seriously, radical reform is in order. For mathematics to be personally meaningful to students, it must be more exploratory; a sense of agency simply can't come from being trained to apply methods and ideas you've been taught. And, for mathematics to be meaningful, it must be more personally relevant. Here I don't mean the superficial relevance of topics drawn from "real life," e.g., discussions of sharing pizza equitably when students are learning fractions.

Many meaningful examples can be drawn from real life, and they can be mathematized. That, in part, is the general issue of "mathematical literacy" (see Burkhardt & Schoenfeld, in preparation). Issues of social justice can and must be mathematized as well. There is a small body of research and resources along these lines (see, e.g., Gutstein 2006, Gutstein & Peterson 2015), and there needs to be much more. But there's more to be considered than mathematizing real world and social justice contexts in classrooms. The challenge is to design ways for students to do that mathematizing in ways that result in their empowerment – the feelings of agency and identity that make it *natural* to see oneself as someone who can approach meaningful problems and make sense of them. What if we thought about organizing curricula with these goals in mind? I think there are possibilities, if only hypothetical for now. In what follows I briefly outline the pie-in the sky version, and then suggest that it isn't impossible.

Imagine a massive research and development project centered around the creation of multiple-days-to-weeks-long units that feature

- (a) potentially meaningful issues to be addressed or resolved;
- (b) a student-centered pedagogy supporting exploration in ways consistent with the development of student agency; and
- (c) scaffolding for teachers that helps them engage with issues (a) and (b) in increasingly powerful ways over time.

Imagine, further, that the units address a broad range of issues, including

- interesting and important mathematical concepts and practices;
- meaningful challenges from the "real world" that can profit from mathematizing; *and*,
- issues of social justice.

And, as long as we're imagining things, imagine building the kinds of professional networks that support teachers in leveraging what they've learned from working with such instructional units.

This vision isn't impossible. Evidence shows that carefully designed instructional materials can result not only in student learning, but in teacher change – at scale. The Formative Assessment Lessons (FALs) developed by the Mathematics Assessment Project (see <https://www.map.mathshell.org/lessons.php>) are 2-3 day units that present students with one or more challenges to address, in exploratory fashion. The teacher support provided in the FALs consists of 20-page lesson plans that structure the explorations and help teachers support the students in those explorations. The lesson plans include descriptions of students' likely misconceptions and ways to address them, while maintaining an ambience of inquiry. Studies of FAL implementation indicate significant student learning gains (Herman et al. 2014) and teacher learning (Research for Action 2015). The fact that there are 20 FALs per grade (in grades 6

through 10) means that it is possible to build 50-60 days of instruction per grade in this mode. That's a third of an academic year. If you can do that, it's possible to build a full year's worth of instruction in similar fashion.

The FALs were constructed to be aligned with the Common Core. What if we were to treat some meaningful real world problems the same way? What if we were to treat some social justice issues the same way? What if we were to craft an entire curriculum with a mix of centrally important mathematics, social justice, and applied units? On the one hand, I think that such materials could make a significant difference – and that a funding agency with a sense of vision could help to make some of this happen. On the other hand, I can imagine the prospect of the first complete social justice unit being caricatured on Fox News and catalyzing the next round of the culture wars. I could say more, but this isn't the place to go into such ideas in depth. My intention here is to plant some seeds for thought. Perhaps some of them can be helped to grow.

Acknowledgement

The author gratefully acknowledges comments on earlier versions of this manuscript from Abraham Arcavi, Brantina Chirinda, Heather Fink, Brian Greer, Markku Hannula, Siqi Huang, Xinyu Wei, and Sandra Zuñiga-Ruiz.

References

- Burkhardt, H., & Schoenfeld, A. H. (2019). Formative assessment in mathematics. In R. Bennett, H. Andrade, & G. Cizek (Eds.), *Handbook of Formative Assessment in the Disciplines* (pp. 35-67). New York: Routledge. ISBN 9781138054363.
<https://www.routledgehandbooks.com/doi/10.4324/9781315166933-3>
- Burkhardt, H., & Schoenfeld, A. H. (in press). Assessment and Mathematical Literacy: A Brief Introduction. *International Encyclopedia of Education, 4th Edition*
- Center for American Progress. (2019). Systemic Inequality: Displacement, Exclusion, and Segregation.
<https://www.americanprogress.org/issues/race/reports/2019/08/07/472617/systemic-inequality-displacement-exclusion-segregation/>
- Common Core State Standards Initiative. (2010). <http://www.corestandards.org/>. See specifically the Common Core State Standards for Mathematics, <http://www.corestandards.org/Math/>
- Egede, L., & Walker, R. (2020) Structural Racism, Social Risk Factors, and Covid-19 — A Dangerous Convergence for Black Americans. *New England Journal of Medicine*, DOI: 10.1056/NEJMp2023616
- Gamow, G. (1947). *One Two Three... Infinity*. New York: Viking.
- Gaylord-Harden, N., Adams-Bass, V., Bogan, E., Francis, L., Scott, J., Seaton, E., and Williams, J. (2020). Addressing Inequities in Education: Considerations for Black Children and Youth in the Era of COVID-19. <https://www.srcd.org/research/addressing-inequities-education-considerations-black-children-and-youth-era-covid-19>
- Gutiérrez, R. (2013). The sociopolitical turn in mathematics education. *Journal for Research in Mathematics Education*, 44(1), 37-68. DOI: [10.5951/jresmetheduc.44.1.0037](https://doi.org/10.5951/jresmetheduc.44.1.0037)
- Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice*. Taylor & Francis.
- Gutstein, E., & Peterson, B. (Eds.). (2015). *Rethinking mathematics: Teaching social justice by the numbers*. Milwaukee, WI: Rethinking Schools.
- Halmos, P. (1980) The heart of mathematics. *American Mathematical Monthly*, 87, 519-524.
<https://doi.org/10.1080/00029890.1980.11995081>
- Herman, J., Epstein, S., Leon, S., La Torre Matrundola, D., Reber, S., & Choi, K. (2014). *Implementation and effects of LDC and MDC in Kentucky districts (CRESST Policy Brief No. 13)*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Herman, R., Boruch, R., Powell, R., Flesichman, S., & Maynard, R. (2006). Overcoming the Challenges: A Response to Alan H. Schoenfeld's "What Doesn't Work". *Educational Researcher*, 35(2), 22-23. DOI: [10.3102/0013189X035002022](https://doi.org/10.3102/0013189X035002022)
- Kozol, J. (1992). *Savage inequalities*. New York: Harper Perennial.
- Lappan, G., & Phillips, E. (2009). A designer speaks. *Educational Designer*, 1(3).
<http://www.educationaldesigner.org/ed/volume1/issue3/article11> (Accessed 7 Aug 2013).

- Martin, D. B. (2009). Researching race in mathematics education. *Teachers College Record*, 111, 295-338.
- Martin, D. B. (2019). Equity, inclusion, and antiblackness in mathematics education. *Race Ethnicity and Education*, 22:4, 459-478, DOI: 10.1080/13613324.2019.1592833
- Mason, J., Burton, L., & Stacey, K. (1982). *Thinking Mathematically*. Addison-Wesley Publishing Limited.
- McKinney de Royston, M., & Vossoughi, S. (2021). Fixating on pandemic “learning loss” undermines the need to transform education. *Truthout*, <https://truthout.org/articles/fixating-on-pandemic-learning-loss-undermines-the-need-to-transform-education/>
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Research Council. (2002). *Scientific research in education*. (R. J. Shavelson & L. Towne, Eds.). Washington, DC: National Academy Press.
- Pólya, G. (1945; 2nd edition, 1957). *How to solve it*. Princeton: Princeton University Press.
- Pólya, G. (1954). *Mathematics and plausible reasoning* (Volume 1, *Induction and analogy in mathematics*; Volume 2, *Patterns of plausible inference*). Princeton: Princeton University Press.
- Pólya, G. (1962, 1965/1981). *Mathematical Discovery* (Volume 1, 1962; Volume 2, 1965). Princeton: Princeton University Press. Combined paperback edition, 1981. New York: Wiley.
- Research for Action. (2015). *MDC’s Influence on Teaching and Learning*. Philadelphia, PA: Author. Retrieved from <https://www.researchforaction.org/publications/mdcs-influence-on-teaching-and-learning/>
- Ridgway, J., Crust, R., Burkhardt, H., Wilcox, S., Fisher, L., & Foster, D. (2000). MARS report on the 2000 tests. San Jose, CA: Mathematics Assessment Collaborative.
- Rothstein, R. (2017). *The color of law*. New York: Liveright.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Schoenfeld, A. H. (1988, Spring) When good teaching leads to bad results: The disasters of well taught mathematics classes. *Educational Psychologist*, 23 (2), 145-166. https://doi.org/10.1207/s15326985ep2302_5
- Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, 20 (4), 338-355. DOI: [10.2307/749440](https://doi.org/10.2307/749440)
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning*, pp. 334-370. New York: MacMillan. <https://doi.org/10.1177/002205741619600202>
- Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, 18(1), 253-286. <https://doi.org/10.1177/0895904803260042>

- Schoenfeld, A. H. (2006). What doesn't work: The challenge and failure of the What Works Clearinghouse to conduct meaningful reviews of studies of mathematics curricula. *Educational Researcher*, 35(2), 13-21. <https://doi.org/10.3102/0013189X035002013>
- Schoenfeld, A. H. (2006). Reply to comments from the What Works Clearinghouse on What Doesn't Work. *Educational Researcher*, 35(2), 23.
- Schoenfeld, A. H. (2008). Problem Solving in The United States, 1970-2008: Research and Theory, Practice and Politics. In G. Törner, A. H. Schoenfeld, & K. Reiss (Eds.). *Problem Solving Around the World – Summing up the State of the Art*. Special issue of the *Zentralblatt für Didaktik der Mathematik*: Issue 1, 2008. DOI 10.1007/s11858-007-0038-z
- Schoenfeld, A. H. (2017). Teaching for robust understanding of essential mathematics. In T. McDougal, (Ed.), *Essential Mathematics for the Next Generation: What and How Students Should Learn* (pp. 104-129). Tokyo, Japan: Tokyo Gakuji University.
- Schoenfeld, A.H. (2020a). Mathematical Practices, In Theory and Practice. *ZDM: Mathematics education*, 52, 1163–1175. <https://doi.org/10.1007/s11858-020-01162-w>
- Schoenfeld, A. H. (2020b) Reframing teacher knowledge: a research and development agenda. *ZDM*, 52(2), 359-376. <https://doi.org/10.1007/s11858-019-01057-5>
- Schoenfeld, A. H. (2021). Reflections on 50 years of Research & Development in Science Education: What Have We Learned? And where might we be going? In Hofstein, A., Arcavi, A.A., Eylon, B., & Yarden (Eds.), (2021). *Research & Development in Science Education: What Have We Learned?* (pp. 387 - 412). Rehovot, Israel: Department of Science Teaching, Weizmann Institute of Science.
- Schoenfeld, A. H. (in press). Why are Learning and Teaching Mathematics so Difficult? In M. Danesi, (ed). *Handbook of Cognitive Mathematics*. New York: Springer Nature.
- Schoenfeld, A.H., & Burkhardt, H. (March 20, 2012). Content Specifications the Summative assessment of the Common Core State Standards for Mathematics. A specifications document for the SMARTER Balanced Assessment Consortium. Downloaded December 9, 2011, from <http://www.k12.wa.us/smarter/>. A revised version is on the SBAC web site at <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwi3w-KK3Jb0AhWqJzQIHZwtCUYQFnoECBcQAQ&url=https%3A%2F%2Fportal.smarterbalanced.org%2Flibrary%2Fen%2Fmathematics-content-specifications.pdf&usq=AOvVaw2zREtzKxXyqM-p0Iqg0oOD>,
- Schoenfeld, A. H., & Pearson, P. D. (2009) The reading and math wars. In G. Sykes, B. Schneider, & D. Plank (Eds.), *Handbook of education policy research* (pp. 560-580). New York: Routledge. <https://doi.org/10.3102/0091732X16658650>
- Schoenfeld, A. H., and the Teaching for Robust Understanding Project. (2018). *An Introduction to the Teaching for Robust Understanding (TRU) Framework*. Berkeley, CA: Graduate School of Education. Retrieved from: <<http://truframework.org>> or <https://truframework.org/>.
- Senk, S. L., & Thompson, D. R. (Eds.) (2002) *Standards-based school mathematics curricula: What are they? What do students learn?* Mahwah, NJ: Erlbaum.
- Urban Insittute (2020). <https://www.urban.org/features/structural-racism-america>

US Department of Education. (2002). Introduction: No child left behind.
<https://www2.ed.gov/nclb/overview/intro/index.html>

US Department of Education. (2003). Identifying and Implementing Educational Practices Supported By Rigorous Evidence: A User Friendly Guide.
https://ies.ed.gov/ncee/pubs/evidence_based/evidence_based.asp

Wilkerson, I. (2020). *Caste*. New York: Random House.