ALAN M. HOFMEISTER, Utah State University

Education Reform in Mathematics: A History Ignored?

Abstract: Nationally, public education prepares to meet increasing federal and state accountability requirements. This article examines the history of math education reform and concludes that many present national trends do not provide substantive, valid alternatives to past failed practices. The evidence documents a failure to apply fundamental research and program-development practices to systematically and progressively improve math instruction for all students.

Forty-six years ago, October 4, 1957, to be exact, Sputnik was launched from Kazakhstan. The consternation in the United States was immediate and pervasive. Politicians and editorialists accused the U.S. educational system of falling behind in math and science. To address these concerns, changes were suggested and investments were made.

In summarizing the immediate post-Sputnik reforms, Shulman (1986) stated:

> The emphasis on beefing up the subject matter was matched with a strong concern for inquiry, discovery, and problem solving, for student-initiated activities and divergent thinking and for ascending the heights of Bloom's taxonomy. The opinion leaders were less concerned with the basics than with the more elevated understandings that are needed to be scientifically literate and competitive. (pp. 11–12)

Twenty years ago, the report, *A Nation at Risk* (National Committee on Excellence in Education [Excellence Committee], 1983), was released. This report documented the failure of post-Sputnik efforts to close the gap in math and science achievement between the United States and other industrialized nations. The report, *A Nation at Risk*, again generated a range of "reforms" and investments to address the continuing adverse achievement comparisons between the United States and other nations. These reforms, like their Sputnik predecessors, were generated in a crisis atmosphere. Most reforms placed an emphasis on immediate, not long-term, returns.

The major concerns that followed *A Nation at Risk* generated the 1989 National Council of Teachers of Mathematics (NCTM) Standards (National Council of Teachers of Mathematics, 1989). These standards addressed curriculum, pedagogy, and assessment. A review of Shulman's (1986) previously listed summary of the changes advocated for the post-Sputnik era appear remarkably similar to the content and intent of the 1989 NCTM standards, presented more than 20 years later as a "reform" vehicle to counter the failed post-Sputnik changes in math instruction.

Reports of adverse academic achievement comparisons with other nations and within the United States did generate one attempt at a long-term, program evaluation activity—a national approach to monitoring student progress. For the past 30 years, the U.S.

Journal of Direct Instruction, Vol. 4, No. 1, pp. 5–11. Address correspondence to Alan Hofmeister at hofa@cc.usu.edu

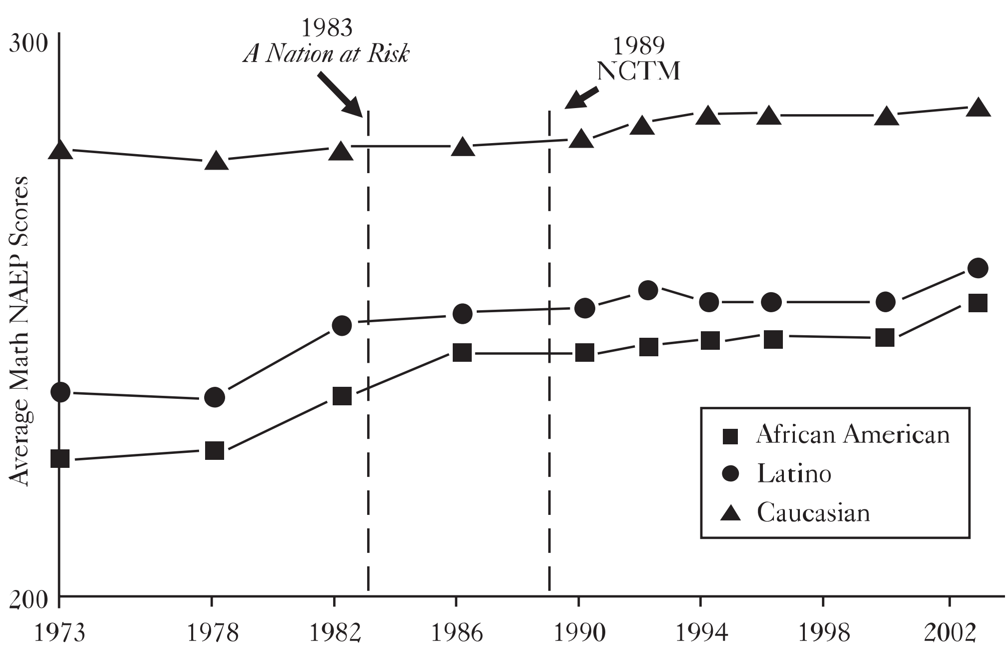
Department of Education, through the National Assessment of Educational Progress (NAEP), has monitored the math knowledge of 9-, 13-, and 17-year-old students. The performance of different subgroups, based on descriptors such as gender and race, has been included. Figure 1 provides a chart of achievement data of 13-year-olds. The data in Figure 1 attempt to document long-term trends in math achievement.

The trends in Figure 1 document an initial narrowing of the gap between Caucasians, African Americans, and Latino students in the post-Sputnik era but not in the post-1989 NCTM standards era. While there have been gains, a recent report (Loveless, 2003) observed that the momentum of these gains seems to be declining. The comparatively modest gains in the long-term NAEP math data trends do not appear to have addressed the original post-Sputnik and *A Nation at Risk* concerns. Our math achievement levels still invoke the same adverse international comparisons as they did 30 years ago (Organization for Economic Cooperation and Development [OECD], 2003).

The reduction in "Cold War" tension has changed, but not removed, national concerns related to academic achievement. Present concerns were summarized in a 1999 National Science Board Report as follows: "In the new global context, a scientifically literate population is vital to the democratic process, a healthy economy, and our quality of life" (p. 3). This report, *Preparing Our Children: Math and Science Education in the National Interest*, further documents the causal links between low achievement and several major societal problems. The

Figure 1

NAEP math scores, 13-year-olds.



needs of national defense, science, and industry were expanded with the notion that, "Educational excellence improves not just the health of science, but everyone's life chances, through productive employment, active citizenship, and continuous learning" (p. 3).

In a break from the rather partisan and competitive relationship among curriculum content areas, the National Science Board (1999) also documented the link between reading failure and societal consequences for the individual and the community. The recognition of the importance of other content, besides math and science, has major implications for K–12 interventions. Too often, the teacher, without guidance, must allocate limited time and resources among ever increasing, conflicting, and competing curriculum demands (Education Trust, 2002).

The initial post-Sputnik emphasis—teaching to the top 20% of students to produce the engineers and scientists—has been replaced by an emphasis on all students, including both high and low achievers. The program descriptor, "No Child Left Behind" (NCLB), now summarizes present priorities. Additionally, NCLB policies are not limited to a single-subject area. Math, reading, and to a lesser extent, science, are included in the present and developing instructional and assessment priorities.

Lessons From the History of Education Reform in Mathematics

Were there lessons to be learned from the post-Sputnik reforms? The answer is, "Yes." Did the changes that followed *A Nation at Risk* and the associated 1989 NCTM standards reflect a benefit from the earlier post-Sputnik experiences? If we examine the NAEP data (see Figure 1) and the international comparative data (OECD, 2003), the answer is clearly, "No."

Despite major investments in instructional interventions, assessment, and research, one lesson is clear: The primary goal of the associated research and development effort was not met. We are not "standing on the shoulders" of those who went before. While the Figure 1 data suggest some modest gains, at least for Caucasian 13-year-olds, these "gains" may be an illusion. The post-1989 NCTM standards era was accompanied by massive federal and state investments. The federal Eisenhower math and science programs dominated K–12 staff development nationwide. Other areas, such as reading instruction, were modestly supported by comparison. Both the advocated, time-intensive NCTM instructional practices, as well as the claim on teacher attention, generated increases in the quantity of time allocated for math instruction. The U.S. Department of Education suggested that at least an 8% increase in math instructional time had been achieved. Such an increase in the quantity of time allocated for math instruction could explain the modest gains in NAEP scores.

A common assumption must be questioned—that changes in the quality of, not the quantity of, instruction caused the achievement gains (Webb, 2002). If there are returns from the post-1989 NCTM era investments, were they cost effective, and did the increased time allocations for math come at the cost of increased reading failure? Such important questions have not been answered despite massive investments by the U.S. Department of Education and the National Science Foundation in the development and evaluation of "standards-based curricula." The NAEP math assessment instruments have "evolved" (and continue to evolve) to further confuse the long-term trends (U.S. Government, 2003).

Recent observations (Vinovskis, 2003) suggest that the math reforms and associated interventions of the past 20 years lacked any pretense at applying the systematic research and development consistent with the level of national concern. The 1989 NCTM standards were implemented in K–12 classrooms and teacher education programs as though they had research validity. Bass (2000) and Howe (1998) suggested that the standards were the product of a rather narrow collaborative process in which educators predominated. Educational researchers, the public, and mathematicians had modest roles. Bass reported that the firm embrace of the NCTM standards by professional education and government may not have been shared by the scientific community and the broad public.

Observers (Bass, 2000; Howe, 1998) have commented on a contradiction between the essential nature of mathematic reasoning and the implementation of NCTM standards in K–12 curricula and in teacher education. Bass stated: "In mathematics the role of proof is to produce conviction or validation in the workplace. Conviction is established by satisfactory performance of a completed product" (p. 11). There is considerable evidence from national and international sources to suggest that K–12 and teacher education investments linked to the 1989 NCTM standards failed the test of workplace validation. For example, in a recent analysis of the history of national education reform, Vinovskis (2003) reported, "They [educators and government] keep spinning their wheels ...Everybody wants their own plan, so we go through these big initiatives in an approach where we try a fad and don't really measure it to see whether it really works" (p. 88).

In April 2000, the revision of the 1989 NCTM standards was published as *Principles and Standards for School Mathematics* (PSSM). Has this revision addressed the major concerns documented in recent syntheses of the history and effectiveness of math education reform (Education Trust, 2002; Ravitch, 2003; Vinovskis, 2003)? The following major concerns remain.

Concern 1: Evidence of Instructional Effectiveness

As states increase math requirements for graduation, the need for refined evaluation practices increases. If courses, such as geometry, move from elective to required and the quantity of time in math instruction increases significantly, how will any gains in achievement be explained? Without more refined evaluation practices that discriminate among qualitative and quantitative factors, achievement gains could occur in spite of, and not because of, the PSSM "instructional quality" recommendations.

Howe (1998), in summarizing the concerns of the American Mathematical Society, stated:

> There has been extensive development of new mathematics programs that attempt to implement the reform (NCTM) ideas. NSF has sponsored several programs at the high school and junior high school levels, and there are also commercially developed new programs. These programs have been a rich mine of controversy. It is remarkably difficult to get a clear picture of effects. There is heartfelt anecdotal testimony both positive and negative. When statistical evidence, favorable or unfavorable, becomes available, its significance can be disputed on the basis of the issues sketched above as well as on technical statistical grounds. (p. 246)

The 1989 NCTM standards did not come from an open-ended, objective synthesis of the research. Later attempts to link to the research were, at best, post hoc procedures that lacked credibility. Longitudinal reading research (National Reading Panel, 2000) moved through an open-ended search for correlative variables and then to studies to identify and validate causal relationships among instructional practices and student outcomes. There appears to be no math counterpart to this 30-year reading research program, and no reasons provided to support this lack of programmatic research.

The 20 years of evaluation and research that followed have not countered concerns that the 1989 standards were seriously flawed and lacked the research validation to support the national adoptions in public schools. Just as troubling, these unvalidated standards are used by the National Council for Accreditation of Teacher Education (NCATE). When the nation's largest teacher education accreditation organization uses NCTM standards, we must question the long-term impact on the next generation of teachers and their students. The 2000 revision of the NCTM standards contains no substantive plans or recommendations to address the lack of research validation.

Concern 2: Instructional Equity

Unlike the 1989 NCTM standards, the 2000 PSSM standards do contain language that recognizes earlier omissions, particularly the neglect of "at-risk" students and "students with exceptional promise." The first principle in the PSSM is the Equity Principle. In a Mathematical Association of America (MAA) review, Ross (2000) reported the presence of language supporting equity but not the presence of substance. Ross stated: "Unfortunately, there is little follow-up on these (equity) statements nor any clear indication as to how the different needs of students should be addressed" (p. 4).

This lack of substance represents a major concern given the trends in Figure 1 and the lack of evidence to suggest we are closing the gap among student populations.

Concern 3: Respect for the Scientific Method, the Research on Instructional Design, and Validation Requirements of Federal and State Laws

The level of research support for the design of instructional practice and curriculum hierarchies in the PSSM remains a major concern. The curriculum is still considered a mile wide and an inch deep (Education Trust, 2000; Ross, 2002). Well-documented past concerns related to the presence of problematic curriculum designs and the lack of reference to the wealth of cognitive science research on problem solving are not addressed (Hofmeister, 1993; Hutchinson, 1993; Mercer, Harris, & Miller, 1993).

The emphasis on vague learning processes at the expense of objective student outcomes remains. Raimi (2001) stated:

> Almost anything in the way of content to be remembered can be omitted from a school mathematics program without running afoul of PSSM, provided the pedagogy is right and the process suitably "exploratory." "Explore," "develop," and "understand," and their variants are much more prominent in the (PSSM) text than "know," "prove," and "remember."(p. 1)

A constant theme in concerns related to NCTM standards is the complex, conflicting, and convoluted formulations of recommendations and the associated theoretical formulations. We have addressed one of the major requirements of a scientific approach that is absent; namely, the process of generating and testing assumptions through correlational and causal phases. Another major scientific requirement, parsimony, appears to be absent. Useful theories should explain "the most with the least," all else being equal. Unnecessarily complex recommendations reduce the transfer of research to practice.

Instructional descriptors such as "systematic," "explicit," and "direct" suggest a parsimonious approach and exemplify the findings from the past 30 years of effective teaching and reading instruction research. The previously referenced descriptors from the NCTM standards appear to be the antithesis of descriptors such as systematic, explicit, and direct (Ross, 2000).

The more recent 2002 NAEP assessments, reflected in Figure 1, indicate an increase, but not a change, in the long-term trend. Efforts to give credit to qualitative reforms in curriculum or pedagogy have little support (Webb, 2002). The confusion is further compounded by the moving goal posts, namely, the NAEP math assessments themselves (U.S. Government, 2003). Issues, such as the limited emphasis on test items assessing operations with numbers, continue long-term, unresolved, post-Sputnik concerns. Questions addressing the range and generalizability of problem-solving skills, and the dependence on the automaticity of preskills in the curriculum hierarchy, exemplify such basic unanswered questions. Any pretense at substance by committee decisions, in lieu of research on student outcomes, raises even more basic questions about educational claims to a client-based profession.

An open letter to the U.S. Secretary of Education published in *The Washington Post* (1999) challenged the validity of government decisions to classify several "NCTM-like" math programs as "exemplary or promising." The letter was generated by 222 mathematicians and scientists, including Nobel Laureates and distinguished faculty from universities including Princeton, Stanford, Cal Tech, and Harvard. These concerns further documented the lack of substantive programmatic efforts to use the tools of science to improve the quality of math instruction systematically and progressively.

Conclusion

In the United States, the public, through elected representatives, establishes laws and sets policies to guide those in service professions, such as medicine and education. These laws are consistently client-referenced. Accountability in these public services requires us to ensure that our professional practices remain subservient to the needs of the client. In education, federal and state accountability laws require us to discriminate between teaching practices and student outcomes. We must progressively and systematically refine our teaching practices based on valid measures of student outcomes. Observers from the National Trust, the Mathematics Association of America, the American Mathematics Society, and the National Academy of Sciences have questioned the commitment to the client. The concerns are based on the comparatively modest level of commitment to produce the evidence of client impact needed to justify the massive investments of public trust and fiscal resources over the past 20 years.

After 20 years and no substantive evidence of success, we need to move quickly and directly to replace unvalidated "standards." The level of institutionalization of these unvalidated "standards" in assessment, in K–12 education, and in teacher education is massive. The effort to turn this around must be focused and equally massive. We must place the needs of all students, and particularly our most vulnerable students, ahead of the organizational and government agency egos that brought us to our present state.

For reading instruction we have a programmatic, longitudinal, research base (Lyon, 2000). We have a highly credible synthesis of this reading research base conducted by the National Reading Panel (2000). This synthesis validates the federal Reading First pedagogical and curriculum mandates that guide K–3 reading instruction nationwide. Where are the validated, national recommendations for school boards, instructional leaders, and teachers searching for practices to address national accountability requirements for math instruction?

References

Bass, H. (2000). *Report: Mathematical preparation of the technical work force.* Washington, DC: The National Academy of Sciences.

Education Trust. (2002, Summer). Add it up: Mathematics education in the U.S. does not compute. *Thinking K–16, 6*(1), 1–23.

Hofmeister, A. M. (1993). Elitism and reform in school mathematics. *Remedial and Special Education, 14*(6), 8–13.

Howe, R. (1998). The AMS and mathematics education: The revision of the "NCTM Standards." *Notices of the AMS, 45*(2), 243–247.

Hutchinson, N. L. (1993). Second invited response: Students with disabilities and mathematics education reform—Let the dialogue begin. *Remedial and Special Education, 14*(6), 20–23.

Loveless, T. (2003). *Event summary: The state of student achievement.* Washington, DC: Brookings Institution.

Lyon, R. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: United States Department of Health and Human Services, National Institute of Public Health, Public Health Service. National Institute of Child Health and Human Development, National Reading Panel, and Government Printing Office.

Mercer, C. D., Harris, C. A., & Miller, S. P. (1993). First invited response: Reforming reforms in mathematics. *Remedial and Special Education, 14*(6), 14–19.

National Committee on Excellence in Education (Excellence Commission). (1983). *A Nation at Risk.* Washington, DC: U.S. Department of Education.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author

National Reading Panel. (2000, April 13). *National Reading Panel summary report.* Washington, DC: National Institute of Child Health and Human Development.

National Science Board. (1999). *Preparing our children: Math and science education in the national interest.* (Rep: NSB Document Reference No. 99-31).

Open letter to Secretary of Education Richard Riley [Paid advertisement]. (1999, November 18). *The Washington Post*.

Organization for Economic Cooperation and Development (OECD). (2003). *Education at a glance*. New York: United Nations Educational, Scientific and Cultural Organization (UNESCO).

Raimi, R. (2001). Standards in school mathematics. *AMS Notices, 48*(2), 236–241.

Ravitch, D. (2003). A historic document. In P. E. Peterson (Ed.), *Our schools and our future: Are we still at risk?* (pp. 25–36). Stanford, CA: Hoover Press.

Ross, K. A. (2000). *The MAA and the new NCTM standards.* (Rep.) Washington, DC: The Mathematical Association of America.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 4–36). New York: Macmillan.

U.S. Government. (2003, October 23). National Assessment Governing Board; Meeting. Federal Register, 68, 205, 60651.

Vinovskis, M. (2003). Missed opportunities: Why the federal response to the report was inadequate. In D. T. Gordon (Ed.), *American education 20 years after A Nation at Risk* (pp. 154–182). Boston: Harvard Press.

Webb, N. L. (2002, April). Comparisons of NAEP mathematics score trends and patterns for SSI and Non-SSI states. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.