Reflections on the National Mathematics Advisory Panel Final Report

Anthony E. Kelly

In March 2008, the National Mathematics Advisory Panel final report was released. This report was produced in response to an executive order from President George W. Bush. The report is important because of its subject matter—improving mathematics teaching and learning—its historically significant genesis, and the strong position that the report takes on the primacy of quantitative methods in education research. The author briefly introduces the report and then draws attention to some of the main points in the commentaries offered in this special issue of *Educational Researcher*. The special issue ends with a rejoinder from the chair and co-chair of the report.

Keywords: federal government; mathematics education; National Mathematics Advisory Panel; policy; research methods

Next to learning to read, developing mathematical competency represents the single largest investment by educational systems worldwide. Even the simplest economic activity is dependent on mathematical skills. Further, the linkage between mathematics and science has led to numerous policy documents, including recent calls to improve international competitiveness (Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology, 2007).

Reflecting the importance of mathematics education, the president of the United States established the National Mathematics Advisory Panel (NMAP) via Executive Order 13398 in April of 2006:

> To help keep America competitive, support American talent and creativity, encourage innovation throughout the American economy, and help State, local, territorial, and tribal governments give the Nation's children and youth the education they need to succeed, it shall be the policy of the United States to foster greater knowledge of and improved performance in mathematics among American students. (NMAP, 2008a, p. 20519)

The goals for the Panel included (a) making evidence-based statements about which mathematics topics should be taught in K–12, (b) specifying how mathematics should be taught and with which materials and curricula, (c) identifying how teachers should be trained and supported professionally, and (d) suggesting directions for future research and establishing standards of quality for this research. Regarding the final point, the Panel noted that

> to produce a steady supply of high-quality research that is relevant to classroom instruction, national capacity must be increased: More researchers in the field of mathematics education must be prepared, venues for research must be made accessible, and a pipeline of research must be funded that extends from the basic science of learning, to the rigorous development of materials and interventions to help improve learning, to field studies in classrooms. The most important criterion for this research is scientific rigor, ensuring trustworthy knowledge in areas of national need. (NMAP, 2008a, p. 65)

The Panel's recommendations are advisory to the president and to U.S. Secretary of Education Margaret Spellings. The report currently stands as the main advisory document for the next White House administration and Congress. Because the report was produced by the U.S. Department of Education, it is likely to prove influential for education policy and may guide the design and implementation of funding programs.

The significant implications of this report for the mathematics education community, and especially for the readership of *Educational Researcher,* include not only the policy recommendations and potential funding implications for mathematics education but also the Panel's adoption of a strict and narrow definition of "scientific evidence" and an almost exclusive endorsement of quantitative methods at the expense of qualitative approaches. Thus, the existence of the report not only affects the curriculum in K–12 education but also raises fundamental questions about the scientific character of education research.

The Panel produced a comprehensive series of reports. In addition to the final report is a set of task group reports on (a) standards of evidence, (b) conceptual knowledge and skills, (c) learning processes, (d) teachers and teacher education, (e) instructional practices, (f) instructional materials, and (g) assessment, as well as (h) a national survey of Algebra I teachers (NMAP, 2008c). The entire report with subreports and summaries is available on the Web (NMAP, 2008b; *National Math Panel,* 2008). The reader is exhorted to read beyond the summary documents, where fine examples of scholarly analyses may be found.

Following the release of the report in March 2008, a campaign to implement the recommendations of the NMAP was launched, including a presentation at the American Educational

Educational Researcher, Vol. 37, No. 9, pp. 561–564
DOI: 10.3102/0013189X08329353

Research Association (AERA) 2008 annual meeting, a dissemination conference in fall 2008, and the distribution of 160,000 pamphlets for parents and for elementary and middle schools across the United States (Cavanagh, 2008). Web-based sources have also appeared on the U.S. Department of Education's (2008) *Doing What Works* website, including multimedia presentations, recommended practices, and templates for working with various implementation constituencies. Similar resources for promoting the "critical foundations of algebra" are slated to appear in January 2009 (Cavanagh, 2008). An interactive symposium on the report will occur at AERA 2009 in San Diego and at the National Council of Teachers of Mathematics national meeting in April 2009 in Washington, D.C.[1]

Commentaries on the Report

The NMAP report has attracted commentary in the U.S. national press, including articles in *Science*, the *New York Times*, and the *Washington Post*. International coverage included Canada's *Globe and Mail*. A critical analysis of the NMAP report appeared in a theme issue of the *Montana Mathematics Enthusiast* (Greer, 2008). A critical review also appeared in the Mathematical Association of America's *MAA Focus* (Ralston, 2008), and *Teachers College Record* (Good, 2008). A more neutral review appeared in *Mathematics Educator* (Moldavan, 2008). The National Council of Teachers of Mathematics (2008) generally endorsed the report's recommendations and called for funding to help in its implementation and for more research to advance the teaching and learning of mathematics.

In addition to these sources, readers may wish to consult related documents such as Lester (2007); Kilpatrick, Swafford, and Findell (2001); the Committee on Programs for Advanced Study of Mathematics and Science in American High Schools (2002); and Shavelson and Towne (2002).

Commentaries in This Issue

This issue features commentaries by a number of researchers who seek to redress what some see as an imbalance in the report. These articles are written from different perspectives: (a) mathematics education research (articles by Boaler; Borko & Whitcomb; Cobb & Jackson; Confrey, Maloney, & Nguyen; Lobato; and Thompson), (b) a more expansive view of the mathematics necessary for lifelong learning (Roschelle, Singleton, Sabelli, Pea, & Bransford), (c) other research genres (Greeno & Collins), (d) a more comprehensive view of the role of randomized clinical trials (Sloane), (e) assessment practices (Shepard), and (e) policy implications (Spillane). A rejoinder to their comments is provided by Benbow and Faulkner (the vice chair and chair of the report, respectively).

The Deconstruction of the NMAP Report

As Spillane notes in his commentary, governments use national panel reports and related white papers as political documents to influence policy, practice, research directions, and funding. National panel reports are thus designed artifacts, rhetorical instruments, whose claims are subject to argumentative analysis (e.g., Kelly & Yin, 2007). The following authors comment on various aspects of the report.

Borko and Whitcomb

As journal editors, Borko and Whitcomb are aware of issues of quality in education research and sympathize with the task that faced the NMAP. However, they disagree with the Panel that randomized clinical trials represent the summative research event and that other methods are merely preparatory for this event. They argue that different research methods have different functions and answer different research questions. Limiting research findings to quantitative studies, they assert, provides a distorted view of the field of mathematics education research, particularly theorizing about the role and impact of teachers. For example, considerations of teachers' mathematical knowledge for teaching in the report were narrow, and the integrated act of teaching and learning was fragmented across task group reports, diluting the construct and potentially distorting the policy implications for teacher education.

Cobb and Jackson

Cobb and Jackson articulate a concern of many of the other commentators that the Panel over-relied on quantitative studies, particularly randomized trials, in defining the quality of mathematics education research. They assert that the imposition of such a filter betrays an ideological position (i.e., experimentalism) that serves to downgrade the importance of context, particularly around issues of equity and measurement (a topic of direct focus by Shepard). They fault the Panel for valuing regularity over process views of causality (citing Maxwell, 2004) and recommend expanding the mathematics topics of interest to include statistics.

Thompson

Thompson faults the Panel for adopting a narrow view of research methods that eliminates from consideration vast amounts of research literature, while retaining for itself a subjective character in its judgments. He agrees with Cobb and Jackson on their characterizing the report as ideological but also sees the choice of research filter as having political overtones.

Boaler

Boaler believes that the authors of the report knowingly created definitional dichotomies (such as between teacher- and student-centered instruction) that undermine scholarship in this area and mislead practitioners. She argues that "gold standard" randomized trials (e.g., Raudenbush, 2005) are unsuited to field settings in education. Boaler shares Holland's aphorism (cited in Briggs, 2008, p. 20) that "a randomized controlled experiment is just a quasi-experiment waiting to happen." She believes that research methods must reflect the phenomenon of learning and teaching mathematics that is realized in messy social settings and that is shaped by, and requires, nontrivial amounts of time, resources, and personnel to enact. By insisting on a filter of strict randomized studies, the Panel makes the ideal the enemy of the good—the good in this case being quasi-experimental and qualitative studies more suited to the realities of complex field research.

Lobato

Lobato faults the Panel on its valuing of regularity over process models of causality. She shows that detailed attention to disciplinary

definitions of mathematics concepts, together with students' misconceptions of these concepts, has clear implications for interpretation of the results of regularity studies. Her analyses contribute directly to Shepard's concerns about the quality of mathematics assessments, which, if in doubt, cast doubt on the dependent measures that form the backbone of the recommendations from randomized trials. Thus the downplaying of qualitative analyses in the report impoverishes its ability to form causal explanations of phenomena (Shadish, Cook, & Campbell, 2002). A key concern for mathematics instruction is not only assigning meaning to symbols but also developing capacity to transfer mathematical ideas to everyday life, to problems in engineering, and to other areas of science. Lobato argues that the report adopts an information-processing model of transfer, which is inadequate to the task of accounting for failures to transfer learning; she proposes alternative models. Lobato also criticizes the Panel's understanding of Vygotsky's theory by noting that core assumptions vary across the theories and are not interchangeable. She refuses to accept the Panel's assumption that the validity of Vygotsky's theory could be tested by fitting it to the assumptions of the behaviorism operative in the use of randomized clinical trials.

Shepard

Shepard criticizes the report on a number of grounds that have significant implications not only for the construct validity of the measures used but also for the validity of large assessments of mathematics learning. Shepard notes the problem of creating a national test in the absence of a national curriculum, a problem she claims is addressed neither by using narrow measures of mathematics skills nor by attempting to test too many topics. She critically reviews the Panel's take on formative assessment. She faults the Panel for not addressing the problem of teaching the test, where instruction targets the indicator, not the curriculum. She points out how difficult it is to write good items and design reliable and valid tests, and she holds that the Panel lacked the requisite expertise in this critical area. Shepard dismisses the Panel's claim that multiple-choice and open-response items are interchangeable and asserts that the form of the test can directly influence what is assessed (for a physics education example, see Mestre, 2000).

Roschelle, Singleton, Sabelli, Pea, and Bransford

Roschelle and his colleagues approach the Panel's report from the perspective of a National Science Foundation–supported Science of Learning Center on learning in formal and informal settings. Roschelle et al. suggest that the report should (a) embrace the use of a broader set of research methods, (b) seek a more expansive view of mathematics as a modeling tool both inside and outside of the classroom, (c) advocate for a theoretically integrated view of resources for improving mathematics learning (including the use of technology, teacher knowledge, assessment practices, etc.), (d) recommend moving away from methods that assess the impact of any one factor in isolation, and (e) call for a better integration of the findings of the Task Group on Learning Processes into the synthetic analysis of the report and its recommendations.

Greeno and Collins

Greeno and Collins also fault the report for its narrow filter for research. Like Thompson, they believe that the judgment as to what constitutes quality research and evidence is best determined by the entire community of researchers in a field, and not by governmental edict. They argue that important research methods that have grown out of mathematics education research itself (e.g., Kelly & Lesh, 2000) and newer methods that have arisen from the learning sciences, such as design research (e.g., Barab, 2004; Kelly, 2004; Kelly, Lesh & Baek, 2008), together with sociocultural (or activity-theoretical, or situative) framings of learning and teaching, cannot be simply dismissed as nonscientific activities. Greeno and Collins cite a National Academy of Education (1999) project as a viable alternative to the Panel's approach to improving mathematics achievement.

Sloane

Sloane is generally sympathetic to the research direction chosen by the Panel. He values the "what works" framing of the Panel. He faults the use of randomized trials of small experimental studies and large field studies as a means for answering "what works" questions on the grounds that the approach is piecemeal. He argues for a more explicit continuum of research projects in a larger portfolio or program of research. He outlines possible phases of research that he believes better map the drug trial model from medicine to the practice of teaching and learning. In the absence of such an articulated phase model, he asserts that the Panel conflated efficacy and effectiveness trials in its call for more experimental research.

Confrey, Maloney, and Nguyen

Drawing on prior experience in writing national reports, Confrey and her colleagues outline six conditions that they claim lead to high-quality reports. They claim that three conditions were violated in preparing the NMAP report: (a) The composition of the panel was unbalanced with respect to the appropriate expertise and thus misrepresented the consensus view of mathematics education research; (b) the report failed to articulate and fairly and consistently apply coherent methodological standards; and (c) the articulation between the task group reports and the NMAP report was inaccurate and lacked intellectual integrity. These flaws, they claim, weakened the report. Confrey et al. end their article with an anecdote of the misuse, as they see it, of the report at the state level in Missouri.

Spillane

Spillane examines the NMAP report from the perspective of policy and implementation. He notes that the report is primarily an exhortative document (one without the force of law) and, as such, depends on the goodwill and commitments of hierarchies of actors for its realization. These actors include not only those at the federal level but also state actors, school policy makers and implementers (including teachers and parents), and a growing set of "extrasystem" actors such as book publishers, testing companies, and tutoring services. Spillane argues that the impact of the report is unlikely to be direct and immediate. Rather, it will be

hindered by differing interpretations of the meaning of its recommendations and their incomplete and inconsistent implementation, by variable application of the use of incentive schemes, and particularly under the current economic conditions, by scarcity of resources.

As the guest editor, I hope that this special issue of *Educational Researcher* will not only present diverse perspectives on substantive research in mathematics education but also contribute to the current discussions of how methodological approaches are matched to and shaped by questions and objectives. The articles herein are intended to broaden the terms of the ongoing discussion of effective instruction as well as to draw sharp distinctions where there is disagreement.

NOTE

[1]Schedules have not been set at press time. Please search programs for "Kelly," "National Mathematics Advisory Panel," or "George Mason University."

REFERENCES

Briggs, D. C. (2008). Comments on Slavin: Synthesizing causal inferences. *Educational Researcher, 37,* 15–22

Cavanagh, S. (2008). Efforts undertaken to publicize math panel's results. *Education Week, 28*(8), 6.

Committee on Programs for Advanced Study of Mathematics and Science in American High Schools (2002). *Learning and understanding: Improving advanced study of mathematics and science in U.S. high schools* Washington, DC: National Academies Press.

Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology. (2007). *Rising above the gathering storm.* Washington, DC: National Academy Press.

Good, T. (2008, April 7). Teaching mathematics in Grades 3–5 classrooms. *Teachers College Record.* Retrieved November 20, 2008, from http://www.tcrecord.org/Content.asp?contentid=15196

Greer, B. (Ed). (2008). Critical notice on the National Mathematics Panel advisory report. *Montana Mathematics Enthusiast, 5*(2–3), 365–428.

Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13,* 115–128.

Kelly, A. E., & Lesh, R. (Eds.). (2000). *Handbook of research design in mathematics and science education.* Mahwah, NJ: Lawrence Erlbaum.

Kelly, A. E., & Yin, R. (2007). Strengthening structured abstracts for educational research. *Educational Researcher, 36,* 133–138.

Kelly, A. E, Lesh, R. A., & Baek, J. Y. (Eds.). (2008). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* New York: Routledge.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academies Press.

Lester, F. (Ed.). (2007). *Second handbook of research on mathematics teaching and learning.* Charlotte, NC: Information Age.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3–11.

Mestre, J. P. (2000). Progress in research: The interplay among theory, research questions, and measurement techniques. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 151–168). Mahwah, NJ: Lawrence Erlbaum.

Moldavan, C. (2008). Ruminations on the final report of the National Mathematics Panel. *Mathematics Educator, 18*(1), 2–7.

National Academy of Education. (1999). *Recommendations regarding research priorities: An advisory report to the National Educational Research Policy and Priorities Board* (A. L. Brown & J. G. Greeno, Co-Chairs). New York: Author.

National Council of Teachers of Mathematics. (2008). *NCTM response to National Mathematics Advisory Panel report.* Retrieved November 11, 2009, from http://www.nctm.org/NMPresponse.aspx

National Math Panel. (2008, March). Retrieved November 11, 2008, from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-factsheet.html

National Mathematics Advisory Panel. (2008a). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: U.S. Department of Education. Retrieved November 11, 2008, from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

National Mathematics Advisory Panel. (2008b). *Foundations for success: Findings and recommendations from the National Mathematics Advisory Panel* [brochure]. Retrieved November 11, 2008, from http://www.ed.gov/about/bdscomm/list/mathpanel/nmp-brochure.pdf

National Mathematics Advisory Panel. (2008c). *Reports of the task groups and subcommittees.* Washington, DC: U.S. Department of Education. Retrieved November 11, 2008, from http://www.ed.gov/about/bdscomm/list/mathpanel/reports.html

Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher, 34*(5), 25–31.

Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education.* Washington, DC: National Academies Press.

Ralston, A. (2008). A nation still at risk. *MAA Focus, 28*(6), 18–19.

U.S. Department of Education. (2008). *Doing what works.* Retrieved November 11, 2008, from http://dww.ed.gov/topic/topic_landing.cfm?PA_ID=8&T_ID=20&Tab=1

AUTHOR

ANTHONY E. KELLY is a professor of educational psychology in the College of Education and Human Development, George Mason University, 4087 University Drive, Fairfax VA 22030, *akelly1@gmu.edu.* His research interests include design research methods, cognitive neuroscience and mathematics learning.

Manuscript received November 7, 2008
Accepted November 12, 2008