Rejoinder to the Critiques of the National Mathematics Advisory Panel Final Report

Camilla Persson Benbow and Larry R. Faulkner

The authors respond to comments and criticisms made of *Foundations for Success: The Final Report of the National Mathematics Advisory Panel* (2008). Given the requirements of the president's executive order, constraints of time, and the need to make recommendations generalizable to national policy, the Panel adopted appropriately high evidentiary standards for its recommendations. The authors also respond to criticisms voiced on Panel findings regarding mathematics curricula, instructional practices, learning processes, teachers and teacher education, and assessment. While scholars should continue to develop promising programs of research, experimental or otherwise, the report offers a solid basis for improving math achievement.

Keywords: assessment; curricula; education policy; instruction; learning; mathematics education; National Mathematics Advisory Panel; research methodology; standards of evidence; teacher education; teachers

To begin, a word of thanks is in order to the many scholars who have devoted considerable time to reading and studying *Foundations for Success: The Final Report of the National Mathematics Advisory Panel* (NMAP; 2008). The critiques they have offered in this issue of *Educational Researcher* amount to the first concerted effort by scholars to analyze the Panel's various recommendations for improving mathematics education in the United States. Their comments represent the development of a dialogue in the research community that adds intellectual depth to what has become a national policy discussion. In particular, Anthony (Eamonn) Kelly is to be thanked for having conceived the idea for this issue and for tapping a number of prominent scholars to weigh in with their criticisms.

The Report in Context

Roschelle, Singleton, Sabelli, Pea, and Bransford ("Mathematics Worth Knowing, Resources Worth Growing, Research Worth Noting," this issue of *Educational Researcher,* pp. 610–617) frame the positive elements of the report successfully and commend its "sharper focus" on mathematical content, the need for stronger mathematics knowledge among teachers, and our efforts to steer clear of some longstanding and unproductive debates. In underlining the report's emphasis on

- the importance of particular topics (e.g., rational numbers);
- the properties of focus, coherence, and closure in mathematics curricula;
- the required integration of concepts and procedures; and
- the need to boost teachers' mathematical knowledge,

the authors indicate that they appreciate not only the breadth of the task but also its particular challenges and its importance.

James P. Spillane's article, "Policy, Politics, and the National Mathematics Advisory Panel Report" (this issue of *Educational Researcher,* pp. 638–644), situates the Panel's work in the context of "ongoing policy discourses and texts." Spillane appears to understand that our report is neither fish nor fowl. Situated in the world of federal politics and governmental policy, it is advisory only in nature. It does not constitute actual policy. And, although it draws on scientific evidence developed in the educational research community, it is not itself a scientific document. Spillane also apprehends the curious situation into which a report like ours falls. He notes growing federal activism in educational areas while being mindful that, to a great extent, educational policy continues to be determined largely at state and district levels and under the influence of many extragovernmental actors. Cobb and Jackson, in "The Consequences of Experimentalism" (this issue of *Educational Researcher,* pp. 573–581), reinforce Spillane's skepticism about how experimental results might be embodied in practices and policies by educational leaders or individual teachers at local levels.

After all, even an activist piece of legislation like No Child Left Behind left it to the states to determine their own benchmarks for adequate yearly progress. Under such circumstances, Spillane's cautions about the likelihood that our suggestions will be implemented, especially given current economic conditions, seem merited. However, this should not deter us from pursuing the focused curricular changes advocated. Too much is at stake.

Standards of Evidence

By far the strongest critique in these articles is of the standards of evidence chosen by the Panel in order to adhere to the president's charge that we provide advice based on "the best available scientific evidence" (Executive Order No. 13398, 2006). Several authors take this opportunity to attack the so-called gold standard

Educational Researcher, Vol. 37, No. 9, pp. 645–648
DOI: 10.3102/0013189X08329195

in educational research. This is misdirected and unfortunate. Of course, it would have been naive of the Panel to assume that any criteria it established for what constitutes "scientific" evidence would not be subject to inspection and debate within the scholarly community. As Greeno and Collins ("Commentary on the Final Report of the National Mathematics Advisory Panel," this issue of *Educational Researcher,* pp. 618–623) rightly point out, "The methods of every scientific field evolve, and one of the most important functions of a scientific community is to continually examine, criticize, and improve its methods" (p. 618).

Boaler, in "When Politics Took the Place of Inquiry" (this issue of *Educational Researcher,* pp. 588–594), is among those authors who take the Panel to task for relying primarily on experimental evidence. This article makes a valuable argument for the role of quasi-experimental and qualitative studies in developing our understanding of mathematics learning over the past 20 years. However, even if we acknowledge that any definition of scientific evidence is at best a moving target, or a partial representation of the methods within a given field, the Panel nonetheless had to establish certain criteria for the admission of studies.

As we state in the report's Executive Summary, we "set a high bar for admitting research results into consideration" (NMAP, 2008, p. xvi). We deliberately favored studies that tested hypotheses that met high methodological standards and that had "been replicated with diverse samples of students under conditions that warrant generalization" (p. 81). At the same time, we accepted the challenge to hear from experts in the field and to speak from our own professional expertise on issues of importance. Task groups were granted further discretion to go where the relevant evidence led them. Borko and Whitcomb ("Teachers, Teaching, and Teacher Education," this issue of *Educational Researcher,* pp. 565–572) express admiration for the Panel's "serious attention to research quality," for its acknowledgment of "a variety of research designs," and its call for more basic research. At the same time, they join others in their concern for the Panel's emphasis on experimental and quasi-experimental evidence, fearing that its endorsement of alternative designs may be lost on the reader.

It should be acknowledged that implicit in the Panel's chosen standards are two challenges. The first is the sheer amount of available research literature. As noted in the report, the Panel reviewed "more than 16,000 research studies and related documents" (NMAP, 2008, p. 82). This is closely related to the challenge of available time. We had only 18 months in which to receive testimony, conduct research (including a national survey of algebra teachers), debate findings, draw conclusions, formulate recommendations, and prepare interim and final reports. Members of the Panel were tasked with doing this while continuing careers with high levels of responsibility. Establishing limits was necessary if we were to serve the president, the secretary of education, and the nation. Bearing this in mind, it is perhaps most useful to see the report not as the end of an initiative but as the first step of a more formalized process that moves from rhetorical handwringing to the framing of initiatives and the development of future research directions.

Volume of research and time, however, were not the only constraining factors on the Panel's work. Our given task was to provide advice to the president and the secretary that could be generalized to policy. This required us to be very careful. Studies that are suggestive or descriptive, or interventions lacking strong methodologies, for all their merits and potential, are difficult to recommend on a national level. Many alternative studies discussed in this issue are examples of ideas the scholarly community may be able to develop further to enrich national policy discussions. As Cobb and Jackson state, "Studies conducted using other methodologies can also produce forms of knowledge that have value in their own right and that can inform experimental studies and complement the forms of knowledge they produce" (p. 576). Finbarr Sloane's article, "Randomized Trials in Mathematics Education" (this issue of *Educational Researcher,* pp. 624–630), offers an intriguing pathway (albeit a "working model") to move from basic research to work that clearly proves causal effects. Sloane's distinction between efficacy and effectiveness is particularly useful.

Instructional Methods

On another front, Boaler argues that "the Panel imposed dangerous dichotomies" (p. 588) in its definitions of two broad teaching methods, either teacher directed or student centered. Indeed, most classrooms may steer something more akin to a middle course between these two types of instruction. Far from imposing a dichotomy, we would argue that the Panel sought to do just the opposite. Our recommendation 23 clearly states:

> All-encompassing recommendations that instruction should be entirely "student centered" or "teacher directed" are not supported by research. If such recommendations exist, they should be rescinded. If they are being considered, they should be avoided. High-quality research does not support the exclusive use of either approach. (NMAP, 2008, p. xxii)

Roschelle et al. seem to have understood the Panel's intent "to dissolve false and regressive dichotomies. For example, the report provides a balanced resolution for the incessant, unproductive debates on teacher-centered versus student-centered instruction and on the need to emphasize algorithms versus the need to emphasize concepts" (p. 610). In fact, our recommendations suggest that both teacher-directed and student-centered approaches have merit for certain purposes and under certain circumstances. As noted in the report, for example, Team Assisted Individualization, a cooperative learning strategy, has been shown to improve computation skills but does not appear to help with conceptual understanding (NMAP, 2008, pp. xxii–xxiii), whereas explicit instruction clearly helps struggling students with word problems as well as with computation (p. xxiii). In other words, there is nothing in the report that explicitly discounts the valuable work of the researchers Boaler cites or of Boaler herself. Promising research that was excluded based on methodological criteria should be further developed to better inform policy makers in the future.

Skills, Meaning, and Mathematical Content

Patrick Thompson's article on curricular content ("On Professional Judgment and the National Mathematics Advisory Panel Report," this issue of *Educational Researcher,* pp. 582–587) joins others in questioning our emphasis on experimental evidence, but Thompson also charges the Task Group on Conceptual Knowledge and Skills with relying too much on professional

judgment and being too selective in that reliance: "The net effect, then, is that chapter 4 of the Panel's report is neither informed by basic research in mathematics education nor grounded in high-quality studies of efficacy and effect" (p. 583). Although this task group may have relied more heavily on judgment than did others, it was hardly operating in a vacuum. State standards, current textbooks, National Assessment of Educational Progress (NAEP) objectives, the American Diploma Project benchmarks, international curricula from countries with high TIMMS performance, the National Council of Teachers of Mathematics *Curriculum Focal Points,* a 2007 American College Testing survey, the National Research Council's 2001 report *Adding It Up,* and the Panel's own survey of hundreds of current algebra teachers all informed the task group's recommendations.

Thompson's larger argument, however, is that the report's recommendations give insufficient attention to the question of meaning, or conceptual understanding, and its appropriate role in a mathematics curriculum. The Panel did, in fact, seek to emphasize the need for meaning, in concert with computational fluency and problem-solving skills (see recommendation 10 in NMAP, 2008, p. xix). Although it may be argued that our recommendations were to some extent corrective (algebra teachers told us that their primary challenge was a lack of fluency and computational skill in their students), we insisted in our Principal Messages on the "mutually reinforcing benefits of conceptual understanding, procedural fluency, and automatic . . . recall of facts" (p. xiv). Thompson's advocacy for a curriculum in which a greater sense of meaning supports "an interest in mathematics that results in taking more, and higher level, coursework" (p. 584) is certainly not in conflict with the spirit of our recommendations.

Similarly, we would continue to stress that our recommendations for the Critical Foundations for Algebra and the Benchmarks for the Critical Foundations should not be taken for a self-contained curriculum. These elements represent the skeleton on which it is up to mathematicians, educators, and policy makers to hang critical organs, interacting systems, and flesh. Cobb and Jackson, to cite just one example, ask for a greater emphasis on statistical reasoning. Their statement that "statistical discourse has become the language of power in the public policy arena" (p. 578) is a compelling argument for doing so.

On Learning Processes

Lobato ("On Learning Processes and the National Mathematics Advisory Panel Report," this issue of *Educational Researcher,* pp. 595–601) expresses a similar concern for the development of conceptual understanding. She argues that the Task Group on Learning Processes tilted toward a regularity view of causality while ignoring process causality, which is better at addressing *why* and *how* questions. (Cobb and Jackson join Lobato in this argument in the context of their own critique of experimental methods, which they develop along pragmatic and historical lines.) Lobato also regrets the task group's missed opportunity to call more attention to the "conceptual understandings that should be achieved in algebra and pre-algebra topics" (p. 596). As noted above, we felt that this had been stressed. Moreover, we would not disagree that Vygotskian and other theoretical approaches are worthy of continued exploration, as is ongoing development of theories regarding transfer.

Roschelle et al. offer an interesting critique here, too, suggesting that the Panel's impact may have been hindered for structural reasons: "The Panel broke into task groups that were largely populated by members with similar knowledge and interests. Unfortunately, the impact of each of these groups on the final report was significantly diluted" (p. 611). A case in point, they say, is that "the work of the Task Group on Learning Processes was not integrated with the core question of what mathematics is worth knowing" (p. 611). Clearly, stronger interactions between psychologists and mathematics researchers could lead to a full flowering of understanding with the greatest benefits among groups that are most disadvantaged. The article by Roschelle et al. offers some of the most provocative ideas, including the development of identity; "transfer and assessment as *preparation for future learning*"; the proper "*integration* of technology, teacher professional development, and curriculum"; and fresh perspectives on what should be measured and across what timeframes (p. 611). Far from condemning the work of the Task Group on Learning Processes, they laud its "high quality and utility" and call special attention to its work on the achievement gap. It should be noted, however, that Cobb and Jackson argue that the Panel did not address disparities in the quality of instruction that many disadvantaged students or English language learners receive.

On Teachers and Teacher Education

Borko and Whitcomb generally agree with the Panel's concern that current means of measuring teacher content knowledge in mathematics are inadequate, relying primarily on proxies rather than direct measurement. However, they argue that the Panel's recommendations on teachers and teacher training give insufficient attention to other things teachers need to know in addition to mathematical content, notably "pedagogical content knowledge." This may be so, given what we do know about the impact a high-quality teacher can have on student learning. The authors also lament missed "opportunities to integrate findings across task groups," notably between learning processes and teacher preparation: "The discussion of teacher knowledge does not highlight how important it is for teachers to know and understand learning processes" (p. 568). Intriguingly, they encourage more research attention to working conditions, amplifying a recommendation that may not have received sufficient emphasis in our final report. With others, they wish for a more "coherent pedagogical vision" in the report, and they point to recent work on core instructional practices as promising.

On Assessment

In her article on assessment, Lorrie Shepard (this issue of *Educational Researcher,* pp. 602–609), raises several questions while acknowledging that the Panel is "at its best . . . when talking about mathematics content" (p. 602). Shepard generally concurs with the need the Panel voiced for a better alignment between NAEP, international standards, and proposed mathematical content.

Perhaps most important, Shepard points out "the dilemmas inherent in creating a national assessment in a country without a national curriculum" (p. 603). This is a thorny problem that the Panel did, in fact, sidestep. There are powerful arguments both

for and against a national curriculum; had we elected to visit this issue, the Panel would have become bogged down in political territory well beyond its purview. As a result, our attempt to put mathematical content foremost would have been too easily discounted by advocates on either side of this issue. Much as with the misleading debate between teacher-directed or student-centered instruction, the Panel sought to steer its ship within the bounds of current realities. This does not mean that these realities will not or should not change.

Finally, Shepard offers frequent reminders that discussions of assessment need to account for the effects, both pro and con, of teaching to the test. She also argues for broader interpretation of the means of formative assessment. In both instances her arguments have merit.

Does the NMAP Report Have a Future?

One article, "Breaching the Conditions for Success for a National Advisory Panel," by Confrey, Mahoney, and Nguyen, calls the report as a whole into question and raises serious questions about its "trustworthiness" and "validity." The authors go on to suggest that the report has already had an undermining effect, offering as evidence objections raised to a "proposed new set of K–12 mathematics curriculum standards" developed by the Missouri Department of Elementary and Secondary Education with the Mathematics, Engineering, Technology and Science Alliance. A letter signed by "39 faculty members, mostly mathematicians, including a former University of Missouri System president" (p. 636) argued that the proposed standards were too narrow and ignored the work of the Panel.

According to Confrey et al., "The flawed NMAP report is contributing to a marginalization of mathematics educators and to the neglect of decades of research on children's learning of mathematics" (p. 636). This is a distressing accusation, if valid. On the other hand, perhaps we should take heart that the report is generating serious discussion at the state level.

Conclusion

The NMAP's 18-month journey from initial inquiry to the production of task group reports, determination of recommendations, and production of its final report was fast paced and at times frenetic. The scholarly critiques offered here by researchers help to situate the report in alternative contexts and to broaden the discussion in constructive ways. Nonetheless, it will be helpful if scholars, too, bear in the mind the constraints of the federal policy-making process, constraints that are aimed at serving students and the nation by demanding evidence of beneficial effects. We all have much to learn. Most important, we all have much to do.

REFERENCES

Boaler, J. (2008). When politics took the place of inquiry: A response to the National Mathematics Advisory Panel's review of instructional practices. *Educational Researcher, 37*(9), 588–594.

Borko, H., & Whitcomb, J. A. (2008). Teachers, teaching, and teacher education: Comments on the National Mathematics Advisory Panel's report. *Educational Researcher, 37*(9), 565–572.

Cobb, P., & Jackson, K. (2008). The consequences of experimentalism in formulating recommendations for policy and practice in mathematics education. *Educational Researcher, 37*(9), 573–581.

Confrey, J., Maloney, A. P., & Nguyen, K. H. (2008). Breaching the conditions for success for a national advisory panel. *Educational Researcher, 37*(9), 631–637.

Executive Order No. 13398. (2006, April 21). *Federal Register, 71*(77), 20519.

Greeno, J. G., & Collins, A. (2008). Commentary on the final report of the National Mathematics Advisory Panel. *Educational Researcher, 37*(9), 618–623.

Lobato, J. (2008). On learning processes and the National Mathematics Advisory Panel report. *Educational Researcher, 37*(9), 595–601.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: U.S. Department of Education.

Roschelle, J., Singleton, C., Sabelli, N., Pea, R., & Bransford, J. D. (2008). Mathematics worth knowing, resources worth growing, research worth noting: A response to the National Mathematics Advisory Panel report. *Educational Researcher, 37*(9), 610–617.

Shepard, L. A. (2008). Commentary on the National Mathematics Advisory Panel recommendations on assessment. *Educational Researcher, 37*(9), 602–609.

Sloane, F. C. (2008). Randomized trials in mathematics education: Recalibrating the proposed high watermark. *Educational Researcher, 37*(9), 624–630.

Spillane, J. P. (2008). Policy, politics, and the National Mathematics Advisory Panel report: Topology, functions, and limits. *Educational Researcher, 37*(9), 638–644.

Thompson, P. W. (2008). On professional judgment and the National Mathematics Advisory Panel report: Curricular content. *Educational Researcher, 37*(9), 582–587.

AUTHORS

CAMILLA PERSSON BENBOW served as vice chair of the National Mathematics Advisory Panel. She is the Patricia and Rodes Hart Dean of Education and Human Development at Vanderbilt University's Peabody College, Peabody No. 329, 230 Appleton Place, Nashville, TN 37203-5721; *camilla.benbow@vanderbilt.edu*. She co-directs the Study of Mathematically Precocious Youth, a longitudinal study examining the developmental trajectories of more than 5,000 individuals throughout the life span.

LARRY R. FAULKNER served as chair of the National Mathematics Advisory Panel. He is the president emeritus of the University of Texas at Austin and president of the Houston Endowment, 600 Travis, Suite 6400, Houston, TX 77002-3000; *LFaulkner@houstonendowment.org*.

Manuscript received November 6, 2008
Accepted November 6, 2008