



When Politics Took the Place of Inquiry: A Response to the National Mathematics Advisory Panel's Review of Instructional Practices

Jo Boaler

The National Mathematics Advisory Panel was given the task of reviewing research on the instructional practices that enable students to learn mathematics successfully. The authors argue that in conducting its review, which appears in *Foundations for Success: The Final Report of the National Mathematics Advisory Panel* (2008), the Panel imposed dangerous dichotomies, opposing extreme forms of teaching against one another. Such dichotomies bear little relation to the realities of mathematics classrooms in the United States and belie the research that has been conducted in mathematics education in the past 20 years. In addition, the methodological restrictions imposed by the Panel rendered the field of mathematics education virtually invisible.

Keywords: experimental research; generalizability; mathematics education; National Mathematics Advisory Panel; qualitative research; research methodology; teacher research; teaching approaches

The National Mathematics Advisory Panel (2008) was asked to consider the timeless question, What instructional practices enable students to learn mathematics most successfully? This is a critical question for our society, but I will argue in this article that the opportunity to advance understanding and dissemination of effective teaching approaches was lost because the Panel

- constructed a dichotomy between two definitions of teaching that bear little relation to the realities of U.S. classrooms;
- imposed definitions of two forms of teaching that were, by its own admission, “extreme” (Gersten et al., 2008, p. 30); and
- employed such methodological bias in the selection of research studies that the field of mathematics education was rendered virtually invisible.

In offering a critique of the Panel's criteria, I discuss the potential of quasi-experimental studies in advancing our understanding of effective approaches, as well as other field-defining studies

in mathematics education that challenge the idea that causal claims can only be made in relation to quantitative research (Maxwell, 2004).

Promoting Misconceptions About Teaching

The National Mathematics Advisory Panel was given a broad question concerning the most effective instructional practices in mathematics. The Panel's task was to find research studies that would inform its question. A subset of the Panel (the Task Group on Instructional Practices) chose to reduce its focus to six smaller questions that the group decided to be the most critical and that led its members to consider research on areas such as technology, assessment, and “gifted students” (Gersten et al., 2008). One of the most critical areas the task group chose to consider was the impact of different teaching approaches on student learning. In framing its review, the task group posed this question: “How effective is teacher-directed instruction in mathematics in comparison to student-centered approaches?” (Gersten et al., 2008, p. 12).

This question, in ways I will expand on below, sets up a dangerous dichotomy between two forms of teaching that bear little relation to the reality of mathematics classrooms in the United States. Indeed the question is reminiscent of the sort of dichotomous thinking that characterized the “math wars”—a series of unproductive and heated exchanges between advocates of different teaching approaches (Becker & Jacobs, 2000; Boaler, 2008b; Schoenfeld, 2004; Wilson, 2003). The arguments that have raged in mathematics parallel those in other subject areas, such as reading and history, and mirror a more general divide between advocates of constructivist and those of more didactic teaching in the United States (Rosen, 2000). In promoting this dichotomy, the Panel revealed a lack of knowledge of the ways the field of mathematics education has progressed in the past 20 years. Fortunately, mathematics education has moved beyond such dichotomized thinking to a broader appreciation of the varied and complex roles in which effective teachers of mathematics need to engage (Boaler, 2003; Chazan & Ball, 1999; Hiebert et al., 1997; Kilpatrick, Swafford, & Findell, 2001; Lobato, Clarke, & Ellis, 2005; Sherin, 2002).

Research in mathematics education has shown, conclusively, that effective teaching of mathematics does not only involve the precise presentation of knowledge; it also involves changing the ways children think, building on their current understandings,

and addressing any prior misconceptions (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Cobb, Wood, Yackel, & Perlwitz, 1992; Davis, Maher, & Noddings, 1990; Franke & Kazemi, 2001; Smith, diSessa, & Roschelle, 1994; Steffe & Cobb, 1988). In recognition of this, one of the main contributions of the field of mathematics education research has been the development of an extensive knowledge base documenting learners' common conceptions and misconceptions in different mathematical domains (Grouws, 1992; Kieran, 1992).

Whereas the first 20 years of research in mathematics education concentrated, to a large degree, on learners' conceptions and the developmental trajectories of mathematical ideas, the past 20 years has seen much greater attention to effective teaching approaches (Kilpatrick et al., 2001). Researchers in mathematics have worked to understand and analyse the different aspects of effective teaching, and few have concerned themselves with trying to show that one extreme approach is better than another. Indeed numerous research studies have shown the varied pedagogical acts, including methods of presenting and working with student thinking, in which effective teachers engage (Adler, 1997; Anthony, 1996; Askew, Brown, Rhodes, Johnson, & Wiliam, 1997; Ball, 1993; Boaler, 2003, 2008a, 2008b; Boaler & Staples, 2008; Gutiérrez, 1999; Henningsen & Stein, 1997; Jaworski, 1994; Lampert, 2001). Lobato et al. (2005) explored the types of "telling" that are productive in a teaching approach consonant with constructivist ideals, such as "initiating" ideas and "eliciting" student thoughts. They state that the distinction between constructivist teaching and teacher telling is unproductive and that we must move away from such false choices and work to understand the "more sophisticated range of pedagogical actions" (p. 131) in which effective teachers engage.

The fact that the Panel chose to present a dichotomy between two types of teaching was unfortunate. A more productive question would have asked what we have learned from research about the qualities and characteristics of effective teaching. But the value of the Panel's work was restricted further by what appears to be ideologically driven definitions of the two forms of teaching it sets out to consider. To illustrate, *teacher-directed instruction* is a term that is generally understood to mean a teacher presenting methods to students who watch, listen, and then practice the methods (Boaler, 1998; Good & Grouws, 1977). By contrast, *student-centered instruction*, although it has received more varied definitions, generally implies an approach in which learners are given opportunities to offer their own ideas and to become actively involved in their learning (Cobb, 1994; Confrey, 1990). The National Mathematics Advisory Panel did not employ the definitions that are understood by experts in the field; instead the Panel imposed its own definitions, saying that the difference between student-centered and teacher-directed instruction is that in the former the students are "doing the teaching of the mathematics" (Gersten et al., 2008, p. 35, line 18). Thus in searching for research studies that considered the two forms of teaching, the Panel simply asked, "Who is doing the teaching—teachers or students?" (p. 35, line 28).

The task group justified its definition of student-centered teaching saying that in such approaches "teachers facilitate, encourage, and coach but do not explicitly instruct by showing

and explaining how things work" (Gersten et al., 2008, p. 35). This claim is made despite the fact that the field has widely available examples of student-centered teaching—at the elementary level (Ball, 1993; Lampert, 1985, 2001; Lampert & Ball, 1998), the middle school level (Boaler & Humphreys, 2005), and the high school level (Chazan, 2000). In these examples, the teachers engage in the key features of student-centered teaching as defined by the group—emphasizing "student responsibility" and acknowledging "students' experiences, prior knowledge, and interests" (Gersten et al., 2008, p. 35); they also spend considerable amounts of time "showing and explaining how things work" (Gersten et al., 2008, p. 35).

The task group's assertion that student-centered teaching involves teachers handing over the teaching of mathematics to the students is remarkable. Indeed the task group itself acknowledged that such definitions are "extreme" (Gersten et al., 2008, p. 30). The task group admits that it found no studies in which students were doing the teaching of mathematics, and it quotes from the National Research Council's report *Adding It Up* (Kilpatrick et al., 2001), summarizing the different and complex forms of teaching in which student-centered teachers engage.

The idea that *showing* and *explaining* would be absent from student-centered teachers' pedagogy suggests that the task group members either held deep misunderstandings about student-centered teaching or that certain members of the group were engaged in a more politicized exercise. Even the strongest advocates of different teaching approaches would probably find the task group's trivialized definitions of teaching to be inadequate, making any subsequent reviews of research redundant.

The task group's advocacy for an "extreme" and unrealistic definition of student-centered teaching was blamed on conversations held with teachers. This is particularly ironic given the Panel's admonition to the readers of its report to recognize only experimental evidence. With a startling departure from this principle, the task group reports that

teachers told the panel that they understand the expectations of administrators in their districts are that they teach exclusively in teacher-directed ways, essentially as it has been defined here. And, other teachers have said that their administrators are critical unless they are teaching in student-centered ways, again as it has been defined here. Thus, this review was undertaken to highlight these distinctions in ways that will hopefully help policymakers and teachers to engage in practice that is evidence based. (Gersten et al., 2008, p. 35)

This claim, which is given great importance in the task group's report, seems to be based on anecdotal evidence and is not supported by research of any kind. Furthermore, it seems highly implausible that administrators would tell their teachers to use student-centered approaches in which they hand over the teaching of content to the students and refrain from any explaining or showing. Far from providing a scientific review that would be helpful to policy makers and teachers, the ideological nature of the task group's report means it is likely to perpetuate myths and fears about nontraditional teaching as well as provide barriers to any advancement of understanding about the complexities of teachers' work. The task group's misleading definition, which belies the important progress made in the field of mathematics

education research in the past 20 years, will undoubtedly improve the quality of discussions that take place in public and policy arenas in the future. It is now incumbent on researchers in mathematics education to correct the serious errors that have been made.

Not only did the review on instructional practices suffer from the Panel's ideological framing of the issues, but it suffered greatly from the Panel's exclusion of most of the research that has been conducted in mathematics education—particularly the research on the effectiveness of different forms of teaching.

Using Research to Inform Questions of Teaching and Learning

The National Mathematics Advisory Panel chose to focus its review on randomized controlled trials. The task group adhered closely to the evidence standards provided by the What Works Clearinghouse (2008) from the Department of Education and the Institute of Education Sciences (Gersten et al., 2008, p. 4). The task group consulted experiments that either randomized groups of students or assigned students to groups according to pretest scores, despite leading researchers and policy makers warning against such a “rigid definition of research quality” (Feuer, Towne, & Shavelson, 2002, p. 4). The decision to use extremely narrow selection criteria might have made sense if researchers in mathematics education actually used randomized controlled trials, but the task group found only eight studies that satisfied its criteria, and most of these had monitored a particular teaching approach used for a few days.

The decision to restrict research to studies that randomized or equated groups caused serious problems simply because education researchers rarely ever assign students to groups in these ways. When comparing teaching approaches to consider which is more effective, random or equal assignment may be thought of as presenting a research ideal. If students are assigned to random or equal groups and given different treatments, and one treatment results in better outcomes, then researchers have a strong case for making causal statements. Experiments such as these have emanated from medical research, and they lend themselves to the controlled conditions of laboratories. However, when researching learning in complicated places such as schools, such models become highly impractical and, some would say, implausible. Tom Cook, one of the most successful users of randomized trials, illustrated the problem well when he set out to study the impact of the Comer School Development Program (Cook, Hunt, & Murphy, 2000, p. 535). Cook et al. set up a randomized trial with 24 schools, half of which were randomly assigned to the control treatment but then found that 5 of the schools dropped out of the study, for various reasons—including new principals taking over and schools not wanting to be in the control groups. Cook et al. describe the ideal of randomization being “vitiated” (p. 544), and other researchers find that randomized trials are impossible because they cannot persuade schools to treat children as experimental subjects and assign them to different conditions, which, of course, is entirely reasonable.

The National Mathematics Advisory Panel chose to constrain its review to randomized controlled trials and other experiments that had assigned students to equal groups. In its review of

different teaching approaches, the task group found eight studies that fulfilled its criteria. In each of the studies, the innovation that was studied lasted for only a few days (in one case, a few weeks), and the researchers found, unsurprisingly, that the innovation made little or no impact. The findings of the Panel illustrate well why randomized controlled studies are so rare (Cook, 2002) and why education researchers use other studies, including quasi-experimental studies, in their place.

Quasi-Experimental Studies

In natural experiments, or quasi-experiments, researchers compare teaching approaches, not by sorting children into control and experimental groups and applying treatments, but by finding schools that use different approaches and studying their effectiveness. The term *natural* should not be taken to suggest that such research is not purposeful and rigorous; it simply means that researchers find different teaching approaches and study them in their natural settings, rather than creating different conditions and studying them. Researchers in mathematics education conduct quasi-experiments to compare the effect of teaching approaches, not by assigning students to random or equated groups but by following students in groups formed by their schools and using statistical methods to control for prior achievement. Such researchers do not assign students to groups, partly because they do not need to and partly because they want to research schools operating under natural conditions. The task group stated that quasi-experiments could be consulted in its review but only if researchers had assigned students to equal groups based on pretest scores. This, again, followed the instructions of the government's What Works Clearinghouse (2008) in its declaration of standards for the validity of research. This meant that the majority of all quasi-experiments used in mathematics education was eliminated from the task group's search.

Researchers in mathematics education do not need to assign students to groups in quasi-experimental studies, taking control of their education, as they can employ statistical methods to control for differences in student characteristics. Using logistic regression analysis, for example, researchers can control for factors such as prior mathematics achievement, gender, and socioeconomic status. It could be argued that researchers cannot control for every variable that may affect a student in a population, but they can control for all those known to be reasonable (Campbell, cited in Kelly & Yin, 2007). The external validity of a study that does not assign students to groups may be weaker than one that does, but this is compensated by the increased ecological validity of a study that examines the natural operating of a school. Thus quasi-experimentalists can study schools and students working in ways that are realistic and achievable by other schools, rather than ways that have been artificially created by researchers.

One example of a quasi-experimental study ignored by the task group was the study conducted by Burris, Heubert, and Levin (2006). Burris and colleagues studied the impact of giving high- and low-achieving students an accelerated mathematics curriculum that had previously only been offered to a small number of advanced students each year. They were able to research the impact of this innovation carefully, as a Long Island district chose to change its policy and place all students in an advanced

mathematics course in eighth grade. The researchers studied six cohorts of students: three cohorts from immediately before the reform and the first three cohorts after the reform. Thus the researchers did not need to randomly assign students to different groups, taking control of their educational pathways, with all of the ethical implications that would imply. The district changed its approach, creating the conditions for a quasi-experiment. The researchers tracked six cohorts, investigating student achievement, using four different measures, and high school course taking. The six cohorts ranged in size from 152 to 181 students. Four hundred and seventy-seven students were in the earlier cohorts that took mathematics in different ability groups, with higher achieving students taking the accelerated curriculum. Five hundred and eight students were in the later cohorts that were all taught the accelerated curriculum in heterogeneous groups.

To statistically control for the important differences between the student cohorts, Burriss and colleagues used logistic regression analysis controlling for previous mathematics achievement, socioeconomic status, and ethnicity. The results of the study were critically important, as the researchers found that when low-achieving students were given access to advanced curriculum and instruction, their achievement increased significantly, as did the probability of their completing advanced mathematics courses. Concern has been expressed that mixing lower and higher achieving students together in the same classes lowers the performance and participation of high-achieving students (Boaler, 2008b), but the researchers found that the participation of low-, middle-, and high-achieving students all increased when students were grouped heterogeneously and all followed an advanced curriculum, as did the participation of minority students and students of low socioeconomic status.

As the National Mathematics Advisory Panel was considering the way in which different instructional approaches affect the performance of low-achieving students, this study was a critical one to consult and report on. But the study did not meet the narrow criteria imposed on the Panel's search. This study provides just one illustration of how the methodological bias employed by the Panel resulted in the neglect of critically important research in mathematics education.

My own research studies also use statistical and experimental methods to investigate the impact of different mathematics teaching approaches (Boaler, 1998, 2006). In a recent mixed method study, we tracked approximately 700 students over 4 years of mathematics teaching, with approximately half the students experiencing traditional mathematics teaching and the other half engaging in "complex instruction" (Boaler, 2006, 2008a, 2008b; Boaler & Staples, 2008). In line with other quasi-experimental studies, we did not impose different approaches on the schools; instead we found schools that were using different approaches, and we studied their impact. To investigate the ways student learning was shaped by the different approaches, we observed more than 600 hours of classroom lessons, we administered questionnaires to students every year, we conducted in-depth interviews with large numbers of students, we interviewed teachers, and we conducted a range of probing assessments as well as analyzing state tests. We collected a wide range of data so that not only do we know which approach resulted in higher achievement,

but we also understand how the different approaches affected students and which variables were the most critical. We did not assign students to different groups, as this would have involved intervening in the schools' work; instead we found schools that were using different approaches and monitored students in their school groups, taking account of their prior achievement, social class, ethnicity, and gender. Our results were statistically significant, with the students who experienced complex instruction, in which teachers used different pedagogical strategies (including showing and explaining), starting at statistically lower levels but ending at statistically higher achievement levels. The students, who were also from low-income, diverse populations, outperformed students from more suburban schools who followed an approach in which teachers only showed and explained (and students practiced). This study was also ignored.

In contrast to the eight studies that the mathematics Panel reviewed, which followed a handful of students through a few days of different instructional methods, we followed 700 students more than 4 years. Burriss and colleagues (2006) followed six cohorts of students over a year. When the Panel chose studies to include in its review, one would expect consideration to be given to the length of time that students spent with different instructional approaches, or even the numbers of students involved. It seems that these critical variables were all subsumed by the exclusive focus on randomized trials. The Panel was left with a handful of studies that had followed an innovation for a few days or less—two of the eight studies considered the impact of an innovation that was used in one single lesson. But this time limitation is an inevitable outcome of randomized trials, as researchers simply cannot persuade teachers to apply a certain treatment to students for years of their education. If researchers do manage to impose such conditions on schools, serious problems usually arise in the fidelity of treatments in the study. Yet we know that changes in student learning are slow and incremental (Stevenson & Stigler, 1994). High-quality research needs to investigate the impact of different approaches conducted over time. When it comes to the use of randomized controlled trials in education, it may simply be the case that "the ideal is being made the enemy of the good" (as noted by Kelly in his introduction to this special issue).

The field of mathematics education has developed an important knowledge base on the impact of different teaching approaches, particularly over the past 20 years when statistical techniques have been at their most advanced. This knowledge base reflects some of the most important virtues of good research in that it is cumulative, often longitudinal, and theoretically informed (Hargreaves, 1999), but it was ignored by the Panel (see, e.g., Balfanz, Mac Iver, & Byrnes, 2006; Boaler, 1998, 2006, 2008a, 2008b; Boaler & Staples, 2008; Cobb et al., 1991, 1992; Maher, 1991; Post et al., 2008; Reys, Reys, Lapan, Holliday, & Wasman, 2003; Riordan & Noyce, 2001; Schoenfeld, 2002; Tarr et al., 2008). The fact that these studies were comparative or experimental does not mean that they were more important than other nonexperimental studies in mathematics education. Rather, they illustrate the narrowness of the criteria used by the Panel to judge good quality research—criteria that were so narrow they did not even include other experimental work.

Qualitative Research

The Panel excluded not only nonrandomized experimental studies but also most of the research that has ever been conducted in mathematics education. Researchers in mathematics education, like researchers in other fields, choose different methods to answer different questions. One of the most influential and important research studies in mathematics education is a single case study of a single child. Erlwanger (1975) provided the field with an analysis of one boy's understanding of mathematics, detailing the way in which "Benny's" instructional approach had caused him to develop a series of misconceptions about mathematics. In essence, Benny had learned to follow mathematical rules without understanding why the rules worked or where they came from. Because Benny had learned that mathematics was all about rule following, he invented a few more rules of his own, some of which resulted in his gaining correct answers, using flawed mathematics. Erlwanger realized that the mathematics approach that Benny was following had caused him to develop an approach to mathematics learning that was deeply problematic. In studying Benny's mathematical work and thinking, Erlwanger uncovered an important link between certain teaching approaches and student behaviors that resonated with mathematics teachers everywhere.

Erlwanger's (1975) detailed and close work with one student contributed a great deal to the field, not because he showed through trials that one approach led to such behaviors and another did not, but because he was able to provide the detail and the texture in his analysis that enabled people to see and understand the link between the teaching approach Benny experienced and the mathematical thinking he developed. There is a common belief that qualitative studies of a particular case are not generalizable because researchers have not looked across multiple settings. I disagree—qualitative case studies *can* provide highly generalizable findings, not by showing something repeated across cases but by providing the depth of observation and analysis that enables readers to understand a connection or phenomenon clearly and judge its applicability to other cases. The degree of generalizability rests not only with the number of cases consulted or the randomization of subjects, but with the power of the observation and analysis produced within a study. Erlwanger's analysis was powerful, and his findings have contributed to an improved understanding in our field.

The issue of generalizability is at the forefront of calls for randomized controlled trials, but many researchers would challenge the idea of a direct link between quantitative data and causal findings. Maxwell (2004), for example, argues that causality can be identified even in single cases and that meaning and interpretive understanding are essential in determining causal explanation. This realization calls into question the hierarchical ordering of research methods, with randomized controlled trials defining the gold standard (U.S. Department of Education, Institute of Education Sciences, & National Center for Education Evaluation and Regional Assistance, 2003, p. iii). Analyses such as Erlwanger's (1975) play a significant role in understanding the causality between certain teaching approaches and student learning, and it is part of a knowledge

base that constitutes mathematics education and that is understood by experts in the field.

Mathematics education, as a field, has developed through contributions from many different types of research study, informed by different disciplinary perspectives (Sierpiska & Kilpatrick, 1998). It is my strong belief that we have reached a point in our field where we can draw from these different studies, looking across them in a purposeful way (Kelly & Yin, 2007) to give clear directions that can help teaching and learning. In a sadly regressive move, the Panel handicapped its own analysis and review by imposing unjustified conceptual and methodological limitations, causing it to ignore most of the research that has been produced in the field.

Conclusion

The National Mathematics Advisory Panel was handed a difficult and important task. Its posing of dichotomized questions, bearing little relation to the realities of classrooms, combined with its false definitions of teaching and its decision to exclude most research in mathematics education meant that the Panel's conclusions were weak and the opportunity it faced—to report and disseminate the advances that have been made in our understanding of effective mathematics teaching—was lost. If the Panel had reviewed the range of evidence that exists on mathematics instructional approaches and student learning, results appearing in the highest quality, peer-reviewed scientific journals, it would have drawn very different conclusions about the effectiveness of different mathematics approaches. The report, as it stands, could have serious negative consequences for public understanding of different forms of mathematics teaching and for student learning.

More seriously perhaps, the National Mathematics Advisory Panel's report presents a case of a government controlling not only the membership of a panel chosen to review research—a panel dominated by educational conservatives rather than mathematics researchers—but the forms of knowledge admissible in the public domain. In its adherence to government directions (Gersten et al., 2008, p. 213), resulting in the disregard of the field of mathematics education research, the Panel's report communicates the view that the government, rather than academic researchers, should decide on the forms of knowledge that are legitimate in our pursuit of understandings about ways to help children learn (U.S. Department of Education et al., 2003). When governments step in to control research and knowledge production, limiting the methods used by researchers and the forms of knowledge acceptable, to the extent that a whole field of research is invalidated, then it is time to acknowledge that America's celebrated freedom—of thought and inquiry—has been dealt a very serious blow.

REFERENCES

- Adler, J. (1997). A participatory-inquiry approach and the mediation of mathematical knowledge in a multilingual classroom. *Educational Studies in Mathematics*, 33(3), 235–258.
- Anthony, G. (1996). Active learning in a constructivist framework. *Educational Studies in Mathematics*, 31(4), 349–369.
- Askew, M., Brown, M., Rhodes, V., Johnson, D., & Wiliam, D. (1997). *Effective teachers of numeracy: Final report*. London: King's College.

- Balfanz, R., Mac Iver, D. J., & Byrnes, V. (2006). The implementation and impact of evidence-based mathematics reforms in high-poverty middle schools: A multi-site, multi-year study. *Journal for Research in Mathematics Education*, 37(1), 33–64
- Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary mathematics. *Elementary School Journal*, 93(4), 373–397.
- Becker, J., & Jacob, B. (2000, March). California school mathematics politics: The anti-reform of 1997–1999. *Phi Delta Kappan*, pp. 529–537.
- Boaler, J. (1998). Open and closed mathematics: Student experiences and understandings. *Journal for Research in Mathematics Education*, 29(1), 41–62.
- Boaler, J. (2003). Studying and capturing the complexity of practice: the case of the dance of agency. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 2003 Joint Meeting of PME and PMENA* (Vol. 1, pp. 3–16). Honolulu, HI: CRDG.
- Boaler, J. (2006). “Opening Our Ideas”: How a detracked mathematics approach promoted respect, responsibility, and high achievement. *Theory Into Practice*, 45(1), 40–46.
- Boaler, J. (2008a). Promoting “relational equity” and high mathematics achievement through an innovative mixed ability approach. *British Educational Research Journal*, 34(2), 167–194.
- Boaler, J. (2008b). *What's math got to do with it? Helping children love their least favorite subject—and why it's important for America*. New York: Viking.
- Boaler, J., & Humphreys, C. (2005). *Connecting mathematical ideas: Middle school video cases to support teaching and learning*. Portsmouth, NH: Heinemann.
- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers' College Record*, 110(3), 608–645.
- Burris, C., Heubert, J. P., & Levin, H. (2006). Accelerating mathematics achievement using heterogeneous grouping. *American Educational Research Journal* 43(1), 103–134.
- Carpenter, T., Fennema, E., Peterson, P., Chiang, C.-P., & Loeff, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499–531.
- Chazan, D. (2000). *Beyond formulas in mathematics and teaching: Dynamics of the high school algebra classroom*. New York: Teachers College Press.
- Chazan, D., & Ball, D. L. (1999). Beyond being told not to tell. *For the Learning of Mathematics*, 9(2), 2–10.
- Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, 23(7), 13–20.
- Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., et al. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education*, 22(1), 3–29.
- Cobb, P., Wood, T., Yackel, E., & Perlwitz, M. (1992). A follow-up assessment of a second-grade problem-centered mathematics project. *Educational Studies in Mathematics*, 23(5), 483–504.
- Confrey, J. (1990). What constructivism implies for teaching. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (pp. 107–124). Reston, VA: NCTM.
- Cook, T. D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Education Evaluation and Policy Analysis*, 24(3), 175–199.
- Cook, T. D., Hunt, H. D., & Murphy, R. F. (2000). Comer's school development program in Chicago: A theory-based evaluation. *American Educational Research Journal*, 37(2), 535–597.
- Davis, R. B., Maher, C. A., & Noddings, N. (1990). Constructivist views on the teaching and learning of mathematics. *Journal for Research in Mathematics Education*, Monograph No. 4, 125–146.
- Erlwanger, S. H. (1975). Case studies of children's conceptions of mathematics: Part 1. *Journal of Children's Mathematical Behaviour*, 1(3), 157–283.
- Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher*, 31(8), 4–14.
- Franke, M., & Kazemi, E. (2001). Learning to teach mathematics: Focus on student thinking. *Theory Into Practice*, 40(2), 102–109.
- Gersten, R., Ferrini-Mundy, J., Benbow, C., Clements, D. H., Loveless, T., Williams, V., et al. (2008). *Chapter 6: Report of the Task Group on Instructional Practices*. Washington, DC: U.S. Department of Education. Retrieved October 26, 2008, <http://www.ed.gov/about/bdscomm/list/mathpanel/report/instructional-practices.pdf>
- Good, T. L., & Grouws, D. A. (1977). Teaching effects: A process-product study in fourth-grade mathematics classrooms. *Journal of Teacher Education*, 28(3), 49–54.
- Grouws, D. A. (Ed.). (1992). *Handbook of research on mathematics teaching and learning*. New York: Macmillan.
- Gutiérrez, R. (1999). Advancing urban Latina/o youth in mathematics: Lessons from an effective high school mathematics department. *Urban Review*, 31(3), 263–281.
- Hargreaves, D. (1999). The knowledge-creating school. *British Journal of Educational Studies*, 47(2), 122–144.
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524–549.
- Hiebert, J., Carpenter, T., Fennema, E., Fuson, K., Wearne, D., Murray, H., et al. (1997). *Making Sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.
- Jaworski, B. (1994). *Investigating mathematics teaching: A constructivist enquiry*. London: Falmer.
- Kelly, A. E., & Yin, R. K. (2007). Strengthening structured abstracts for educational research: The need for claim-based structured abstracts. *Educational Researcher*, 36(3), 133–138.
- Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390–419). New York: Macmillan.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
- Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review*, 55(2), 178–194.
- Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.
- Lampert, M., & Ball, D. (1998). *Teaching, multimedia, and mathematics: Investigations of real practice*. New York: Teachers College Press.
- Lobato, J., Clarke, D., & Ellis, A. B. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education*, 36(2), 101–136.
- Maher, C. (1991). Is dealing with mathematics as a thoughtful subject compatible with maintaining satisfactory test scores? A nine-year study. *Journal of Mathematical Behavior*, 10, 225–248.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33(2), 3–11.

- National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.
- Post, T. R., Harwell, M. R., Davis, J. D., Maeda, Y., Cutler, A., Andersen, E., et al. (2008). Standards-based mathematics curricula and middle-grades students' performance on standardized achievement tests. *Journal for Research in Mathematics Education*, 39(2), 184–212.
- Reys, R., Reys, B., Lapan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education*, 34(1), 74–95.
- Riordan, J., & Noyce, P. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, 32(4), 368–398.
- Rosen, L. (2000). *Calculating concerns: The politics of representation in California's "Mathematics Wars."* San Diego: University of California.
- Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher*, 31(1), 13–25.
- Schoenfeld, A. H. (2004). The mathematics wars. *Educational Policy*, 18(1), 253–286.
- Sherin, M. G. (2002). A balancing act: Developing a discourse community in a mathematics classroom. *Journal of Mathematics Teacher Education*, 5(3), 205–233.
- Sierpiska, A., & Kilpatrick, J. (Eds.). (1998). *Mathematics education as a research domain: A search for identity: An ICMI study*. Dordrecht, Netherlands: Kluwer.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115–163.
- Stevenson, H., & Stigler, J. (1994). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Simon & Schuster.
- Steffe, L. P., & Cobb, P. (1988). *Construction of arithmetical meanings and strategies*. New York: Springer-Verlag.
- Tarr, J. E., Reys, R. E., Reys, B. J., Chávez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, 39(3), 247–280.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Washington, DC: U.S. Department of Education.
- What Works Clearinghouse. (2008, May). *WWC evidence standards for reviewing studies*. Retrieved October 26, 2008, from <http://ies.ed.gov/ncee/wwc/references/iDocViewer/Doc.aspx?docId=2&tocId=3>
- Wilson, S. (2003). *California dreaming: Reforming mathematics education*. New Haven, CT: Yale University Press.

AUTHOR

JO BOALER is the Marie Curie Professor of Mathematics Education at the University of Sussex, Essex House, Falmer, BN1 9QQ, England; jo.boaler@sussex.ac.uk. Her research interests include mathematics teaching and learning, equity, and ability grouping.

Manuscript received July 17, 2008
 Revision received October 15, 2008
 Accepted October 17, 2008