Research in Education: Rigor and Relevance

Samuel Otten

Michigan State University

Division of Science & Mathematics Education

Comprehensive Examination

Fall 2010
Research in Education: Rigor and Relevance

In many academic fields, there are debates with regard to the roles of pure science and applied science as well as the respective value of each endeavor (Stokes, 1997). In science, the sharp distinction is often captured in the separation of science, the search for generalizable knowledge about the world, from technology and engineering, the use of scientific knowledge for practical ends. In mathematics, the separation is also significant with individual mathematicians as well as university departments often feeling the need to attach “pure” or “applied” to their names. (A more extreme example of this separation is the case of statistics, which is considered by some to be applied mathematics but by others, mathematicians and statisticians alike, as a separate domain from mathematics altogether. This complete separation, ironically, makes room for the distinction between pure statistics and applied statistics.) A personification of these tensions was G. H. Hardy, a prominent mathematician in the first half of the 20th Century. He famously stated his hope that none of his mathematical works would ever be put to practical use (Hardy, 1967).

Although its history can be measured in decades rather than centuries, the educational research community has followed in these footsteps of tension (Lagemann, 2002) and, to an extent, division (Shulman, Golde, Bueschel, & Garabedian, 2006) of pure from applied endeavors. What is the impetus behind this seemingly persistent desire to draw a clear line between pure and applied research? What are rationales for upholding the division and what are rationales for dissolving it? One key aspect of this debate, particularly acute in the case of education research, is the tension between the rigor of research and the relevance of research to people’s lives. In the next section, I attempt to
clarify the definitions of and connections between pure and applied science on the one hand and rigorous and relevant research on the other. I then survey arguments that have been put forth in the educational research community with regard to these distinctions before closing with a statement of my own position on the issue.

**Definitions and Connections**

The notion of “pure” science carries with it a connotation of being untainted or uncontaminated. This leads naturally to the question: What is the possible contaminant? The answer, of course, is “application.” Thus, inherent to the framing of pure and applied science is a mutual exclusivity because even a hint of application removes the purity of an endeavor. In an effort to recast the situation, government funding agencies in the late 1980s began using the term “basic research,” which they defined as scientific contributions to “general knowledge and an understanding of nature and its laws” (Stokes, 1997, p. 3). Applied research, then, can be viewed as contributions to practical knowledge or to the solution of a practical problem. With respect to the purposes of research, basic research has the goal of increasing understanding whereas applied research has the goal of improving the human condition. From this perspective, the two are not *a priori* incompatible because it is possible for a scientific endeavor to have multiple goals and multiple outcomes.

These distinctions are intimately related to issues of rigor and relevance in research. Rigor refers in general to the intellectual quality of a study and the extent to which the claims of the study are supported by the research design and analysis. However, the evaluation of rigor is not an objective process but is instead determined within the context of “a set of norms enforced by the community of researchers that shape scientific
understanding” (National Research Council [NRC], 2002, p. 2). Thus, rigor is socially defined within a context of scientific inquiry; in other words, rigor is whatever it is conceived to be by a particular community of scholars and stakeholders\(^1\) at a particular point in time.

Relevance has to do with the extent to which a study addresses a problem faced by a particular group of people or the extent to which it is practically applicable in the world. For example, a study that took place in a bilingual mathematics classroom to determine why the Spanish-speaking students were falling behind their English-speaking counterparts and to provide the teacher with support and possible remedies would be highly relevant for that teacher and those students. Note that relevance is defined with respect to a particular situation, not some notion of general relevance to all; that is, the study above is relevant because it directly addresses a problem in a particular situation—it is not required to apply to all classrooms or even other bilingual classrooms.

Furthermore, note that relevance and rigor are independent dimensions. A study can be both rigorous and relevant (i.e., a high-quality study that is useful in a particular situation), neither rigorous nor relevant (i.e., a low-quality study that is not useful to anyone), rigorous but not relevant (i.e., a high-quality study dealing with a topic that is not pertinent to anyone’s life), or relevant but not rigorous (i.e., a low-quality study that nevertheless helps individuals in a particular situation).

\(^1\) I include stakeholders because it is not only the researchers themselves who have a voice on what is and is not rigorous within their field. In mathematics education, for example, government officials inserted themselves centrally into the discussion of rigor through such things as *What Works Clearinghouse* (2010) and the *Final Report of the National Mathematics Advisory Panel* (2008). In these materials, rigorous research is defined according to several “evidence standards” such as randomization, low attrition, and equivalence of comparison groups. Although these standards may be agreed upon in some communities of education research (e.g., among experimental quantitative researchers), they ignore what it would mean to be rigorous in other communities (e.g., design experiments, qualitative case studies).
Coming back to issues of basic and applied research, we can now place these ideas into dialogue with the notions of rigor and relevance. If basic research is oriented toward producing “general knowledge,” then a prerequisite for such work’s success would be intellectual rigor. For knowledge to be accepted generally the research community would have to be approving of the method and the legitimacy of the claims that produced said knowledge. In this way, basic research entails the notion of rigor. This does not, however, imply the converse. It is not the case that rigor only applies to basic research; applied research can also meet standards of quality of design, validity of reasoning, and reliability of results. As described in the opening paragraphs, the communities of basic researchers and applied researchers are often distinct, and since rigor is defined within a particular intellectual community, the standards of rigor for basic research and for applied research may not be the same, but there can be rigor in both, nonetheless. Therefore, a strict identification of rigor with basic research would be an oversimplification of the issue.

Similarly, it would be fallacious to assume that relevance is strictly a characteristic of applied research. Although it is true that, as stated above, applied research is typically explicitly concerned with addressing practical problems and making a difference in people’s lives, this does not preclude the possibility of basic research being highly relevant. Wigner (1960) wrote a famous essay that documented the many instances of mathematical ideas, developed with no regard for practical application, impacting the sciences and people’s everyday lives in profound ways. (Now, with the rise of the information age, one could easily add several new sections to that essay.) In mathematics education, the project known as Cognitively Guided Instruction (Carpenter,
1986; Carpenter, Fennema, & Franke, 1996; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989) can be viewed as a theory of students’ thinking, evolved from basic research efforts, becoming highly relevant (and highly effective) with respect to classroom teachers’ actual instructional practice through mere exposure to the theory (i.e., the teachers were not told how to implement the theory so there was no real translation of theory into applicable terms). Hence, relevance is not reserved for applied studies but, like rigor, can be a characteristic of both basic and applied research.

Having established the fact that there is not a one-to-one mapping from basic and applied research to the notions of rigor and relevance, the question remains of the relationship between basic and applied research and between rigor and relevance. As noted in the introduction, there is a history and tendency to cleanly separate one from the other. This has been described as a “sharp divide between educational research and scholarship and the practice of education in schools and other settings” (NRC, 2002, p. 14). In the next section, I review a few scholars’ perspectives with regard to this issue.

Conflict or Compatibility?

It is tempting, and for many years it was the custom (Stokes, 1997), to place basic and applied research (or rigor and relevance) at opposite ends of a continuum. One is either working to establish generalizable truths about the universe, free from particularities, or one was working to address a practical problem or improve a particular situation, with no concern for generalizability, and never the twain shall meet. Stokes (1997), however, viewed the distinction from a two-dimensional perspective and identified various quadrants of research (see Table 1, which is highly related to the comment about “multiple goals” above). In the upper-left quadrant are researchers who
are working to increase fundamental understandings but are not concerned with applications of their work. In the lower-right quadrant are researchers who are motivated by applications without concern for increasing fundamental understanding. The lower-left quadrant is empty because this would be essentially pointless research, but the significance of Stokes’ perspective comes in the form of the fourth quadrant—known as Pasteur’s quadrant. Here is where research takes place that simultaneously increases fundamental understanding and has important applications to everyday life. It is worth noting that this quadrant is not empty; the prototypical example is Pasteur’s work on micro-organisms in food and beverages and his development of vaccines were highly applicable yet also pushed forward our understanding in these scientific areas.

Table 1

<table>
<thead>
<tr>
<th>Quest for Fundamental Understanding?</th>
<th>Considerations of Use?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
<td>Pure basic research (e.g., Bohr, Hardy)</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Use-inspired basic research (e.g., Pasteur)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pure applied research (e.g., Edison)</td>
</tr>
</tbody>
</table>

Stokes’ (1997) conceptualization of the relationship between basic and applied research shows that he saw compatibility between the two endeavors. Although he recognized research efforts that were distinct in their aims, he posited the existence of research that is simultaneously basic and applicable. In particular, Stokes viewed this simultaneity as being possible even for a single study.

Another way to view the compatibility of basic and applied research is not at the level of a single study but at the level of a community of research. Lagemann (2008), for
example, argued that the field of education research needs efforts being made in three areas: problem-finding research, problem-solving research, and translational research. Problem-finding has to do with framing and articulating good research questions, which necessarily involves the development and use of theory. Problem-solving is related to the identification and illumination of educational problems as they exist in particular contexts, that is, the “real world.” Translational work, then, has to do with transforming “the findings of research into tools that practitioners and policy makers need” (p. 424). An important point, which will return in the concluding section, is that Lagemann is not calling for an individual researcher or a particular project to necessarily undertake all three of these, but for all of them to be adequately addressed in the educational research community overall.

If Stokes (1997) represents an overlap model of the relationship between basic and applied research, then Lagemann (2008) can be viewed as a bridge model. In particular, she is calling for work in theoretical (i.e., problem-finding) and actual (i.e., problem-solving) domains as well as work that bridges them (i.e., translational), resulting in applications of research for various aspects of practice. This bridge model is quite popular in the mathematics education community, as evidenced by many of the practitioner-oriented journals containing regular sections on leveraging “research into practice” (e.g., Foote, 2008; Rubel, 2010).

Although the bridge may exist, this does not imply that it is well traveled or that both endpoints of the bridge are equally developed. The National Council of Teachers of Mathematics [NCTM] research committee (2006) identified the need to make more frequent use of this bridge and also pointed out the value in its bidirectionality—not only
should research be made more useful and available to teachers, but researchers should also “learn more from the knowledge and insights of practitioners” (p. 76). This sentiment was echoed even more recently in NCTM’s research agenda report (Arbaugh et al., 2010) which reaffirms the linking of research and practice as a priority for the mathematics education community and presents a list of important practice-based questions that are worthy of research attention. One could view this report as an attempt to encourage use-inspired basic research (Stokes, 1997) or as a call to problem-find, problem-solve, and translate (Lagemann, 2008) research for practitioners.

Inherent in the notion of a bridge is the fact that basic research and applied research do have important differences—they are in several ways separate land masses, so to speak. Shulman and his colleagues (2006) recognized this when they called for a separation between advanced degrees in basic research (i.e., Ph.D.) and advanced degrees that are more oriented toward practice (i.e., Ed.D.). Labaree (2003) argued that substantially different worldviews exist between research and practice in education: the former is marked by analysis, universality, and theory, whereas the latter is marked by personal connections and particular experiences. Bulterman-Bos (2008) made a similar point in stating that “both [researchers and teachers] have their own language and their own incentive structures” (p. 412). If, as was discussed above, scholarly rigor is determined by the “worldview” and the “language” and “structures” of an academic discipline, then these differences between the land of research and the land of practice would seem to preclude rigorous research from being practically relevant.

However, education research is inherently practical because it is an effort to build understanding of processes and phenomena central to people’s lives (in contrast with, say,
research on a celestial body in Andromeda). Even the most basic work in this field is likely to have a degree of relevance higher than other fields (such as astronomy). As Biesta and Burbules (2003) phrased it:

> It is widely expected that educational research should generate knowledge that is relevant for the day-to-day practice of educators. Educators do not simply want to know how the world "out there" is. They want knowledge that can inform their actions and activities. The same is true for educational policymakers and politicians. They also seek knowledge that can support and guide their decision making. Educational research, one might say, is not so much research about education as it is research for education. (p. 1, emphasis in original)

So, although rigorous research is often not directly relevant in the world of practice, it is an inbuilt charge of education research to seek this relevance. Dewey (1929) recognized this dilemma when he wrote that “[n]o conclusion of scientific research can be converted into an immediate rule of educational art” (p. 19, quoted in Clements, 2007). Indeed, his identification of education as an art rather than a science anticipated what Labaree and others wrote about as differences in worldview between teachers and researchers. And although Dewey valued highly the methods and rigor of scientific inquiry (Biesta & Burbules, 2003), he also recognized difficulties of relevance with regard to educational practice because each practical situation involves levels of complexity and multitudes of factors that are impossible to account for in a scientific study (Dewey, 1929). He concluded that research could not be transformed directly into “rules of action” for educators but did, nevertheless, have “practical utility” (p. 19, quoted in Clements, 2007)
as an informative agent for educators—what is required as an intermediary step is something akin to Lagemann’s translational research.

Before proceeding to the final section in which I speak more personally about these issues, it is necessary to include another metaphor in addition to the overlap of Pasteur’s quadrant and the bridge between basic research and application, and that is the metaphor of a border crossing. Silver (2003) articulated this metaphor by stating that theory is the currency of research while practical application is the currency of practice:

Although the residents on each side of the border between research and practice have different currency valuation schemes, they can productively engage in exchange. Researchers have much to offer, including theoretical perspectives that might be useful in framing and describing practical issues and problems, research methods that might illustrate data-collection practices with practical utility, and findings that possess sufficient generalizability to support appropriate use in applied settings. Practitioners also have much to offer, including a set of important issues and concerns that could and should be addressed in research, a collection of insights gained in and through practice, and a passionate concern for the improvement of education. The two groups have much to gain from collaboration in the borderlands between research and practice. (p. 183)

Whereas a bridge functions only as a means of passage between two domains, these “borderlands” form a space in their own right where productive collaboration can take place. Rather than research crossing over to practice or practice crossing over to research, the two can meet in the borderlands and exchange currencies with one another. Based on Silver’s description, the exchanges that take place would be highly relevant to both sides.
With respect to rigor, since it is defined by a community and thus defined differently on each side of the border crossing, the borderlands would be a place to negotiate the terms of rigor amongst that joint community. Indeed, this would most likely be one of the key aspects of the “exchange” wherein all of the issues discussed above would be involved.

Conclusion

In my personal estimation, it will be difficult to perfectly balance rigor and relevance in education research as long as we model ourselves on the physical sciences (NRC, 2002; Lagemann, 2002). Part of this “physics envy” is a mimicry of their community’s standards of rigor, such as experimental design with randomization (Institute of Education Sciences, 2010). This standard of rigor does not seem to account for the fact that education and physics deal with substantially different phenomena and thus decreases the likelihood that rigorous educational research of this kind will be relevant to the daily, nuanced lives of teachers and students. One common response to this dilemma, especially since the “social turn” in education research in the late 20th Century (Lerman, 2000), has been to turn sharply away from a physical science conception of rigor in favor of more qualitative, contextually-based approaches where rigor is defined by the attention to detail, the careful consideration of context, and the resonance of results (e.g., Creswell, 2009; Erickson, 1986; Merriam, 1988). This approach, however, while perhaps increasing relevance of the work for the actual participants of the research and those who readily identify with them, is perhaps less relevant to practitioners who do not identify with the participants and seems to contribute to the failure of educational research to accumulate valuable and agreed-upon knowledge (NRC, 2002).
So while I believe research oriented both to the physical sciences and to more interpretative approaches can be intellectually rigorous, neither seems to be the answer with respect to generating a body of knowledge that is relevant for practitioners—the former because it is too context-free and the latter because it is too context-dependent and non-accumulating. Furthermore, my fear is that researchers attempting to individually bridge the basic-applied gap, perhaps aiming for a successful existence in Pasteur’s quadrant, will often fail to make significant contributions to either side. It is already challenging enough to do high quality research in a single camp, so attempting it in two camps simultaneously is, I fear, overly ambitious (though not impossible, I admit).

For me, the balance comes in thinking beyond the level of an individual researcher or a single study. As a community working together, we can divide the labor and strike a more productive balance than each of us can individually. Everyone can strive for rigor with respect to their particular intellectual undertaking, but relevance can be measured as a community rather than at an individual level. Some scholars may choose for a period of time to focus exclusively on the development of theory, with little concern given to practical relevance. Other scholars (or the same scholars at a later date) may focus on using that theory to explain in a deep fashion the workings of a particular classroom while still others may employ the theory in widespread quantitative testing of students. Another group may work to translate the theory into practical tools and may study their implementation. Again, rigor should be the goal in each individual situation, and even though relevance varies from endeavor to endeavor, the entire enterprise would be highly relevant to practice.
This conception of the balance between rigor and relevance, between basic and applied research, has clear connections to the ideas reviewed above (particularly the notions from Lagemann, 2008, of problem-finding, problem-solving, and translational research being spread throughout the community) and also rejects the idea that scientifically-generated knowledge is not relevant to practice (Gross, Levitt, & Lewis, 1997, cited in NRC, 2002; Schwandt, 2002, cited in Clements, 2007) and the idea that quantitative methods are the only valid approach to research (Finn, 2001, cited in NRC, 2002). The difficulty, however, is coordinating the various efforts which can only happen successfully if high levels of respect exist between all involved parties. The theoreticians must respect the translational work that is done with teachers and the qualitative researchers must respect relatively context-free quasi-experimental studies in order for accumulated knowledge and relevance to result. This means that various types of rigor must be recognized, various types of currencies must be exchanged. In fact, the most important borderlands may not be between researchers and practitioners but between the various camps in research itself. We must arrive at an effective and amenable process of exchange amongst ourselves so that the worth of our collective currency may rise—otherwise, we risk having nothing of value to offer in the border crossing with practice.
References


Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119-161). New York, NY: Macmillan.


