UNITY AND DIVERSITY IN EDUCATIONAL RESEARCH: ON HETEROGENEOUS COMMUNITY

Kurt Stemhagen, Virginia Commonwealth University
Bryan R. Warnick, The Ohio State University

The debate rages on between those who endorse scientific standards for educational research—scientifically based research (SBR)—and those who see such a characterization of the research endeavor as unduly narrow and rigid. One unresolved question in this debate involves the place of inquiry that does not aim to produce generalizable knowledge, but instead insists on questioning knowledge claims. The persistence of this question reflects a deep division in the educational research community.

In this paper, we argue for an uneasy reconciliation. We seek to blur the distinction between research that aims at practical knowledge and research that aims at finding limits of knowledge. Exploring limitations is an essential part of making something useful, we argue, while attempting to be practical will always reveal to us our limitations. First, we will present the competing paradigms of "research as generalization and simplification" and "research as exploring the limits of knowledge" as a point of departure and examine the ancient nature of this debate. Next, we take on D.C. Phillips' challenge for philosophers to "take an interest in empirical educational research" and we present two specific instances of educational research in light of this reframing.[1] Finally, we consider the implications of this reconciliation for public policy.

THE CONTEMPORARY DEBATE

How should we compare the value of educational research that seeks simplicity and generalization with that which questions the received wisdom and makes things more complex? The argument for simplification and generalization from the SBR camp is often a practical one. In the face of educators who need to make decisions about what to do, we need knowledge that helps to solve practical problems. As Grover Whitehurst argues, "The people on the front lines of education do not want research minutia, or post-modern musings, or philosophy, or theory, or advocacy, or opinions from educational researchers."[2] The world is complex, to be sure, and it is indeed

---

[1] D.C. Phillips, "The Contested Nature of Empirical Educational Research," *Journal of Philosophy of Education* 39, no. 4 (2005): 578.
[2] Grover J. Whitehurst, "The Institute of Education Sciences: New wine, new bottles," Presidential Invited Session at the annual AERA conference, Los Angeles, CA. http://ies.ed.gov/director/speeches2003/04_22/2003_04_22b.asp.

difficult to know if we can ever find the "truth" about the way the world works. However, people do need to make decisions. The complex world needs to be simplified to be useful. Data need to be summarized; generalizations need to be made. Broad standards can be used to promote broad utility. The SBR researchers, sometimes influenced by post-positivistic empiricism, often do admit that knowledge claims are limited; but they think that we need to do our best in spite of these limitations to produce knowledge that is generally useful.[3]

On the other side of the debate are those who often (but not always) come from the humanities and qualitative research community.[4]   These researchers have often been heavily influenced by post-modernism and post-structuralism, and they value research that stresses the ultimate limitations of knowledge claims. This camp is "likely to be tentative (hesitant) about making broad application of [research] findings because realities are multiple and different."[5]   Generalizations are simulations of reality that at best leave out important factors and at worst promote the agenda of powerful oppressors. A simplified history, for example, can promote the agendas of connected and powerful groups. Simplification and generalization fundamentally distort our understandings of the phenomena studied. Thus, this second group studies the "possibilities, limits, usefulness, and dangers of any theoretical position— including their own—in the production of knowledge and lives."[6]   For these researchers, it is vital to cast a skeptical eye on knowledge claims.

We are thus presented with two competing visions of intellectual inquiry. For one group, the purpose of intellectual inquiry is to produce knowledge that will help educators find improved solutions to pressing practical problems. For the other group, the purpose of inquiry is to deflate our pretensions to knowledge—to show the limits of our knowledge claims. One group tries to answer practical questions, the other tries to question practical answers. In the end, we believe that the fissure in educational research is not so much about qualitative versus quantitative research or standards versus no standards; it is about the competing and equally compelling values of practical *utility* and epistemic *humility*.

The dichotomy that is often drawn between the utility and humility, we want to argue, is unhelpful. Of course, we are not the first to argue for the

---

[3] See, for example, the NRC report's introduction. National Research Council. *Scientific Research in Education* (Washington, DC: National Academy Press, 2002).
[4] The division of values between practicality and humility does not fall along lines of methodological difference. Some quantitative researchers may play the role of Socratic gadfly, while other qualitative researchers may attempt to make their studies more generalizable.
[5] Yvonna S. Lincoln and Egon G. Guba. *Naturalistic Inquiry* (London: Sage, 1985), 42.
[6] Elizabeth Adams St. Pierre, "'Science' Rejects Postmodernism," *Educational Researcher* 31, no. 8, (2002): 25-27.

fruitlessness of such dualisms. Others have also attempted to provide nuanced accounts of the various research typologies and their purposes.[7] What we offer here is a reframing of the current debate in a new language that might clarify and highlight the necessity (and fundability!) of both dispositions of inquiry within heterogeneous research communities.

PLATO: UTILITY AND HUMILITY IN DIALOGIC COMMUNITY

Some may find an air of familiarity in framing the debate in terms of usefulness versus skepticism. This is, in some ways, the debate that raged within the Platonic dialogues. The central protagonist in many of Plato's dialogues is, of course, Socrates. The main antagonists of Socrates are the Sophists—people like Protagoras, Gorgias, and so forth. In Plato, Socrates is the character obsessed with the limitations of knowledge claims, particularly in the so-called early dialogues. The impulse to destroy the inflated knowledge claims of "the wise" proved irresistible. For Socrates, the limitations of knowledge stem from our inability to grasp the world conceptually—we can be shown to never really know what we are talking about. In the later dialogues, this Socratic humility is enacted across the dialogues rather than explicitly stated within them. The Socrates of the middle and late dialogues seems to answer questions rather than question answers. The theories that Plato seems to endorse, however, are continually under attack across the dialogues themselves. The political philosophy of the *Laws*, for example, is different from that of the *Republic*. In this way, the later dialogues, while each may propose individual answers, collectively retain their emphasis on the limitations of knowledge. The dialogic quest continues in a different form.

The Sophists represent the other side of the debate. The Sophists recognized the limitations of knowledge claims. Protagoras, for his part, was a famous agnostic: "Concerning the gods, I have no means of knowing whether they exist or not or of what sort they may be, because of the obscurity of the subject, and the brevity of human life."[8] Seeing the vigorous disagreement among cultures and even among the natural scientists of the day over even the most basic questions, the Sophists became suspicious of the idea that we could know the true nature of things. Knowledge, they admitted, was based more on tradition, convention, and yes, power, rather than on anything more absolute. "Man is the measure of all things," Protagoras famously taught, showing that he was fully aware of the problems with universally applicable truths or standards.[9]

What the Sophists decided to do was largely bracket out this epistemological uncertainty and complexity in the face of pressing practical

---

[7] Again, see the efforts detailed in the NRC report.
[8] Quoted in http://www.iep.utm.edu/p/protagor.htm.
[9] See Plato's *Theaetetus* (152a).

problems. We cannot know the true nature of things, they said, but we can work to solve practical problems within particular social contexts. The Sophists' research questions often dealt with action. Protagoras, for example, asked questions of legal responsibility, for example, was an athlete's death to be attributed to the javelin itself, to the man who threw it, or to the authorities responsible for the conduct of the games?[10] Given this disposition, it makes sense that the focus of the Sophists' teaching was on rhetoric, the most useful art of the day, in the service of civic excellence. People could not know the true nature of things, they admitted, but we can achieve momentary agreements on particular courses of action.

In the Platonic dialogues, these two dispositions of inquiry—this "absurd pair" as Socrates calls their dyad—are placed in opposition.[11] Those who believe that we need useful knowledge in the face of practical problems—the Sophists—are continually assailed by Socrates, who harps on epistemic limitations. Perhaps this is why the Platonic dialogues usually end so oddly. Socrates destroys the arguments of those who claim to know about, say, piety and virtue, but few ever seem to change their minds. The characters cling to concepts that are essential to practical life, even though they cannot fully defend them.

This debate is very similar to what we are encountering in educational research. The Sophists, like the SBR champions, recognize complexity but believe we are justified in trying to simplify and generalize to meet pressing social problems. Socrates, like the critics of SBR, is unwilling to sacrifice complexity for practical utility. The utility of the Sophists is at war with the humility of Socrates. (They all would have felt right at home in the hotly contested AERA session.)

Plato, though, not only points to the tension between these values, but also shows how they can coexist. One way of looking at Plato's dialogues is as an attempt to negotiate the competing values of usefulness and humility. Plato, who lived in a culture facing deep practical problems, seems to have understood that the inclusion of both impulses was essential to a robust conception of human inquiry. Of course, Plato had great respect for the epistemic humility of Socrates, but it is clear that the Socratic character is not always triumphant. Socrates sometimes changes his argument in response to practical objections (see *Republi*c II and the "city of pigs" objection) and he even ends up losing the debate in some of the dialogues (e.g., the *Parmenides*). Rather than simply celebrating Socrates, Plato brings the humble Socrates into debate with those who demand practical answers.

---

[10] This story is related in Plutarch's "Pericles" from the *Lives*. See
http://www.gutenberg.org/dirs/etext96/plivs10.txt.
[11] See *Protagoras* (361a).

The dialogues are the creation of a dialogic community where humility and practical necessity exist in an endless give-and-take. Rarely do Plato's dialogic characters come to agreement; the point is not agreement, the point is co-equal tormentation. Perhaps most important is the fact that Plato did not see the epistemic gadfly as existing outside of the research community looking in, nor did he see practicality as irrelevant to proper intellectual humility. Finding limitations is an essential part of making something useful, while attempting to be practical will always reveal to us our limitations, albeit in sometimes unexpected ways. Plato suggests that usefulness and humility are inter-dependent.

REFRAMING THE CURRENT DEBATE

Plato faced the same question we face in education today: how can we reconcile these competing impulses of intellectual inquiry? How can we recognize the need for useful standards, but also the need for those who question the standards? Plato's dialogues suggest the value of both impulses for any viable research program. Those who accept shared standards of research cannot by themselves constitute a community of inquiry; rather, standards must coexist with those who work to interrogate the standards.

Blurring the distinction between utility-minded researchers trying to solve specific practical problems and humility-oriented researchers seeking to question existing knowledge is justified and important for several reasons. For one thing, the history of social science, particularly educational science, has demonstrated that without epistemic humility, even the most practically-oriented research has often failed to produce sustained utility.[12] A primary reason for this is that research with a strong practical orientation often focuses on comparing methods to arrive at pre-established ends. Epistemic humility, though, pushes us to acknowledge that any "knowledge" gained in research is not useful in-and-of-itself and can only be meaningfully evaluated in light of broader social contexts and purposes.

To be truly useful, research testing the effectiveness of particular methods needs to be sensitive to more than just the aims implied in a singular study. One particular method may be effective in promoting one narrow educational aim, for example, but it might make other aims more difficult to achieve. A study may ask whether a particular method of teaching reading is effective and, for simplicity's sake, look to only one measure of effectiveness (say, phonetic decoding). But in such a study, the wider aims of teaching reading would be ignored and other goals of literacy impeded (focusing on phonetic decoding might, in fact, make learning to love reading more difficult).

---

[12] See Ellen Cliff Lagemann, *An Elusive Science: The Troubling History of Education Research* (Chicago: University Of Chicago Press, 2002).

Since studies that simplify phenomena necessarily leave things out, then critical appraisals ought to be considered an essential part of the research community if research is to be at all useful. There is an epistemic humility in acknowledging that any research finding cannot be assessed independent of the wider social aims and contexts within which the inquiry is rooted.

Balancing epistemic humility with practical utility will require rethinking the means-ends and methods-aims relationships. To be of practical utility, researchers often focus on a narrow view of preexisting educational aims to find more effective methods of achieving those ends. In this sort of inquiry, ends are necessarily outside of inquiry. This ignores, however, the ways in which means and ends need to work in concert. John Dewey, for his part, argued that the sharp separation of means and ends suggested by a view of educational research that focuses narrowly on "what works" is ineffective, even nonsensical. Dewey sought to blur the means-ends distinction, claiming: "Consequently, ends arise and function within action. They are not, as current theories too often imply things lying beyond activity at which the latter is directed."[13]   Dewey explained that means without ends are meaningless; indeed, means without ends are not means at all—they are just senseless activity.

While it is wrong for this reason to focus only on "what works" to solve pre-established practical problems, it is equally wrong to ignore entirely questions of what works. The value of ends can be judged by looking at the practicality of the means. If an aim cannot realistically be reached, it is meaningless as a guide to practical action. For example, if we aim to tailor learning to the radical uniqueness of each child in every possible way, it is likely to prove impossible to achieve—no amount of resources would allow for it. Such considerations of practical method should encourage us to modify somehow that particular end. Taking practicality seriously, then, opens up new ways of criticizing ends.

It is also the case that trying to make generalized claims often reveals new varieties of intellectual humility that we might not have considered before—as physicists tried to extend theories like general relativity and quantum mechanics, they found that both theories break down when both applied to tiny objects of immense mass. Trying to generalize, in this way, pointed to the limits of our knowledge.

Making a sharp separation between research seeking to provide knowledge of practical means and research seeking to question our knowledge (and to question the aims at which our means are directed) only serves to impoverish our community of inquiry. In what follows, we describe two studies in order to explore the interrelations between epistemic humility and

---

[13] John Dewey, *Human Nature and Conduct* (New York: Modern Library, 1922), 223.

practical utility and to show problems that emerge when humility and utility are cordoned off from each other. The possibilities for research to be meaningful increase when methods and values are allowed to commingle under the umbrella of research and when questioning what we can know is allowed to exist *within* the enterprise of trying to solve practical problems—that is, when Socrates exists in dialogue with the Sophist.[14]

<center>COMPARING *PRESCHOOL IN THREE CULTURES* AND *TIMSS*</center>

Preschool in Three Cultures describes a research project in which preschools in Japan, China, and the Untied States were videotaped.[15]   Each video was shown to the teachers and students of the particular preschool, to preschool parents and school personnel in a number of other cities within the country of origin, and to audiences from the other countries. At each stage of video viewing, audiences provided written responses to the video and they also participated in recorded discussions. The results of the study shed light on the effectiveness *and* purposes of particular methods within each class. Perhaps most interesting, it becomes clear that without a broad understanding of how participants from each country conceive of the purposes of preschool as a social institution there would be no way to determine the immediate utility of particular methods. Furthermore, the study problematizes our assumptions regarding what methods are effective and what ends the methods should serve.

Two examples from the study serve to demonstrate these points. First, consider that Japanese viewers bemoaned the sorry situation in American preschools whereby the pupil-teacher ratio was so *low*. With such individual attention from the teacher, they thought that children would have difficulty learning to become part of the group—some even claimed that to learn how to be part of the group is to become human. Such differences become even clearer as one watches the video of the American school with its emphasis on independence and individual expression. We come to better understand what it means to be a member of Japanese society as opposed to being a member of American society. Through this study, something as self-evident and unquestioned as the merits of a low pupil-teacher ratio becomes less obvious. With different social ends, various means take on different values.

The second example from this study involves an oft-misbehaving Japanese boy named Hiroki. Chinese and American audiences were critical of how infrequently Japanese teachers intervened when Hiroki's misbehavior disrupted activities and disturbed his classmates. While American and Chinese audiences worried about Hiroki's effect on other students, Japanese audiences

---

[14] For an interesting description of value's appropriate role in research, see Harry Brighouse, "Is 'Evidence' Enough?," *Education Week*, September 27 (2006): 31.

[15] *Preschool in Three Cultures*, VHS, Joseph Tobin, David, Wu, and Dana Davidson (New Haven, CT: Yale University Press, 1989).

saw his presence as an opportunity for children to learn how to become part of a group. When Japanese teachers were told that American educators wondered if Hiroki might be gifted and that his misbehavior might stem from boredom, their initial reaction was confusion as to what it means to be "gifted." Once the concept was explained, they tended to disagree with this label and to describe him as "about average." Revealingly, one Japanese teacher claimed, "If he's so smart how come he doesn't know how to behave?"[16] Thus, we learn about cultural difference, about divergent views of intelligence, and about the ways in which effective evaluation of particular teaching methods needs to be sensitive to such cultural factors. This is a very practical lesson, indeed.

Gert Biesta claims that a robust understandings of educational research "can perform a technical and a cultural role and that both can have very real practical consequences."[17] *Preschool in Three Cultures* demonstrates this point. The study sheds light on some technical aspects of preschool pedagogy and is helpful as a means of cultural understanding. Moreover, preschool teachers and curriculum makers could, upon exposure to the study, adjust methods and curriculum after reflecting on whether their practices are in tune with the aims of preschool and its place in wider society. Note that the study's practical worth involved ways in which the technical and cultural *interrelate*. It is just this co-mingling of the technical (practical) and the questioning that makes the study so worthwhile. It has utility because of its epistemic humility.

Compare this to the *Trends in International Mathematics and Science Study* (TIMSS).   TIMSS is a sprawling research effort coordinated internationally.[18]   This comparative project has collected mathematics and science achievement data since 1995 in dozens of countries. A random sample of international classrooms are visited in each country and achievement tests are administered, compiled, and analyzed. TIMSS is often held up as a way to judge the comparative quality of mathematics and science education in the United States. The test is somewhat controversial, with much of the criticism directed at finding ways to explain the Untied States' very average performance. Many such criticisms are directed at micro-limitations of the research methodology.   Take 2003 TIMSS itself, in which the only acknowledgement of its potential problems are three very brief sentences simply defining sampling and nonsampling errors, with no further application. The creators of the study apparently see the limits of their statistical procedures and data collection techniques as the only potential problems with their study.

---

[16] Ibid.

[17] Gert Biesta, "Why What Works Won't Work," *Educational Theory*, 57, no. 1, (2007): 21.

[18] Patrick Gonzales. *Highlights from the Trends in International Mathematics and Science Study 2003*. (Washington, DC: U.S. Government Printing Office, 2004).

Holliday and Holliday come closer to getting at the complexity involved in making meaning of inter-cultural comparisons, be they qualitative or quantitative, noting that a "much more important hurdle to overcome is the unique set of cultural factors situated in each country, such as differential national languages, social norms, cultural prides, ethical standards, political systems, educational goals, and school curricula."[19]   Their discussion of culture, however, is restricted to the important but incomplete idea that language translations and differences in pride and motivation render the comparisons suspect.

While it is true that international quantitative comparisons of mathematics and science achievement are difficult for a number of logistical reasons, the fact is that even if the test makers and data analyzers could control for questionable variables or ensure that the sample group and test administration was consistent, problems would remain.   As *Preschool* demonstrates, means and aims are necessarily related and it is likely that different social aims for schools in the various countries suggests that performance according to this one "extra-cultural" metric says little about how a given school system is functioning with respect to its culturally specific social purposes.[20]

Aside from the problems associated with small random sample sizes and other methodological questions, TIMSS loses a valuable opportunity to consider the complexity and limits of knowledge. Recall how in *Preschool*, rather than only seeking objectivity or generalization, the researchers chose to focus on questioning practices and taken-as-given knowledge.   This Socratic humility was what actually made the study useful: through questioning, researchers were able to shed light on important dimensions of preschool teaching and learning that otherwise would have remained hidden.   With TIMSS, in its narrow quest for utility, epistemic humility was left out and, consequently, its results are disappointing in terms of its generation of practical knowledge.   This is not to say that TIMSS should not try to simplify and generalize; rather, it should recognize that such practices are valuable in part because they alert us to what we do not know. As it reported its findings, TIMSS could have started a discussion about the possible relationship between

---

[19] William Holliday and Berchie Holliday, "Why Using International Comparative Math and Science Achievement Data from TIMSS Is Not Helpful," *Educational Forum*, 67, no. 3 (2003): 252.

[20] Berliner and Biddle made exactly this point of earlier international comparisons. See *The Manufactured Crisis* (New York: Addison-Wesley, 1995), 52-53.

schools and society or the ends of education, but did not. Its attempt at being useful could have promoted epistemic humility.

<center>SOCRATES AND RESEARCH TODAY</center>

Philosophers of education and other social foundations faculty are often looked to as the people who worry about aims, while the remaining researchers focus on questions of means and methods. For example, the NRC report, while explaining that aims do matter in education—and that philosophy, history and literary studies contribute to education—seems to suggest that such studies are independent of practical knowledge.[21] In the end, we see this fissure as unfortunate as it is ubiquitous; an integrated model of educational research would allow for the research enterprise to benefit from what community and dialogue have to offer.

Biesta's recent conceptualization of educational research provides a good starting point for the creation of such social structures. His criticisms of the SBR movement hinge on his belief that it is fundamentally antidemocratic: "The problems with evidence-based education, therefore, is not only that it is not sufficiently aware of the role of norms and values in educational decision making; the problem is that it also limits opportunities for educational professionals to exert their judgment about what is educationally desirable in particular situations."[22] We would emphasize that when research is defined so narrowly as not to include the knowledge questioners and their consideration of what values are at stake, then antidemocratic research also tends toward the mindless. And, although it is not equivalent in vehemence or power, the same antidemocratic tendency is sometimes manifest from the other side when the knowledge questioners disparage those who attempt to meet the practical realities that push us toward simplification and generalization. We have tried to point to how antidemocratic divisions along these lines work against Plato's idea of bringing together the "absurd pair"—a heterogeneous community of inquiry.

Recognizing the necessity of the heterogeneous community has implications for public policy as it relates to educational research. For example, there has been much discussion in policy circles about the desirability of the "gold standard" of educational research projects. By using designations like "the gold standard," policy makers valorize large-scale experimental projects with random samples. If our analysis is correct, though, it is impossible to reduce science to the desirability (or not) of research projects taken in isolation. A project, by itself, can never quality qualify as "good science." Good science is a product of communities or large-scale research

---

[21] Richard J. Shavelson and Lisa Towne, eds., *Scientific Research in Education* (Washington, D.C.: National Research Council, National Academy Press, 2002).
[22] Biesta, "Why What Works Won't Work," 20.

programs, not of individual studies. The talk of gold standards when applied to specific projects takes an overly individualistic view of the scientific enterprise.

Instead of valorizing individual projects or methods as being "useful" or "practical," public policy should focus on strengthening the vitality of heterogeneous communities. The specific allocations of resources, of course, are open to debate. But, even though much of the funding for research might go to standards-based research focused on overt practicality, there still needs to be space for funded research that emphasizes intellectual humility. Thus, there should be funding reserved for various projects, conferences, and symposia that reflect the Socratic side of the scientific community. Only by preserving this balance in the research community can we ultimately see that utility promotes humility, and humility promotes utility.