Introduction

What Counts as Evidence and Equity

ALLAN LUKE
Queensland University of Technology

JUDITH GREEN
University of California, Santa Barbara

GREGORY KELLY
Pennsylvania State University

The most durable and robust problem facing education research since mid–20th century is the persistence of educational inequality. Under new economic, technological, and cultural conditions, increasingly diverse populations and communities are facing persistent and emergent patterns of educational inclusion and exclusion. How we name and describe, document and understand educational equality and inequality, inclusion and exclusion, centrality and marginality, then, is *the* issue facing educational systems in economically hard times. In this volume, we bring together reviews on different approaches to the formation and use of evidence in educational policy and reform. We asked the authors to examine different approaches to evidence and to focus on how each reflects particular ways of defining, explaining and framing inequality and equality in educational policy and practice. They also provide critical analyses of evidence-based policies to date, discussing policy assumptions about and impacts on educational equity.

This volume is a companion to the last four volumes of *Review of Research in Education*. In our two previous edited volumes, we focused on "what counts as learning" (Vol. 30), and "what counts as knowledge" (Vol. 32). There we and our editorial boards and authors attempted to broaden and expand debates over policy and practice that were becoming narrowly circumscribed in the transnational push for accountability-driven, evidence-based policy. Our goal in these volumes was to bring state-of-the-art research and theoretical perspectives across traditions and disciplines

in order to identify factors that support and constrain what is possible educationally in and out of classrooms, school, community, and regional policy contexts. In these volumes, we recognized the roles of qualitative, mixed-method, action and practitioner-based research in generating valuable, local and situated issues and opportunities around learning in reform contexts. Adjacent volumes of *Review of Research in Education* have focused on reviews of critical research and development work on race, difference, and diversity (Parker, 2007), and on "at risk" student and youth communities (Gadsen, Artiles, & Davis, 2009). These have expanded our engagement with standpoints and perspectives of those student communities that have experienced educational exclusion and marginalization.

In this volume, authors engage with different angles, frames, and levels for discussing matters of the analytic scale and scope of educational reform (see Bryk & Gomez, 2007): a key response to problems of equity in a two-decade-long push for "evidence-based" social and educational policy in advanced industrial nations and economies. Following the lead of transnational organizations such as the Organization of Economic Cooperation and Development (OECD) and the World Bank, the press for accountability-driven educational policy has extended to emergent economies and countries still building basic school and postsecondary educational infrastructure (see Wiseman, Chapter 1). One result has been the driving of the reform of school and classroom practice on the basis of a single major performance indicator: student achievement on high stakes standardized testing. The laudable goal of many of these policies has been closure of what is now termed *the equity gap*—the differential between general population norms and the performance of identified equity groups: for example, African Americans, Latinos, migrants, second language speakers, and students from low socioeconomic backgrounds (see Lucas & Beresford, Chapter 2; Jordan, Chapter 5).

Not surprisingly, the demand for evidence has opened a Pandora's Box of arguments over the appropriate grounds for documenting and analyzing student socioeconomic, cultural and linguistic background, student performance and achievement, systemic delivery of resources, school-type and structure, and school and teacher practices. It is axiomatic that any policy "fix" or strategic approach is contingent on how the problems, target populations, variable contexts, and factors are defined and parceled out and observed, represented and measured, and analyzed. How we define and describe the contexts and impacts of difference and diversity, then, remains on the policy table. In the United States, Canada, the United Kingdom, Australia and New Zealand, European, and Asian systems, there are robust debates over what kinds of evidence can and should be enlisted to analyze systems' equity performance and to shape and implement innovation and reform.

Since its inception more than a century ago, modern education research has been based on distinctive and, at times, contending conceptions of evidence and, indeed, educational science. These range across descriptive and interpretive, quantitative and qualitative, empirical and hermeneutic approaches that draw from diverse theoretical models of education and schooling, knowledge and culture, the learner and

society (Green, Camilli, & Elmore, 2006). As the sciences *and* arts of education have evolved, differing research paradigms have defined educational phenomena in distinctive terms and provided contending forms of evidence for policymakers, educational system bureaucrats, curriculum developers, teachers, students, and parents.[1] But this cannot be attributed solely to academic paradigm wars, contending sociocultural value systems or, in some cases, media and public misrepresentation of research. It raises foundational questions of what kinds of science and philosophy can and should inform educational practice.

On its face, it is a simple, straightforward idea: that educational and social policy should be based on scientific evidence rather than specific political view, philosophy or social ideology. The use of science to inform the formation of laws, institutions and governance of the secular state has a long and, at times, undistinguished history. It featured prominently in 17th- and 18th-century debates over the rational foundations of the modern nation state, with varied bids for governance to be based on the logics and evidence of the sciences of economics, eugenics, psychology, sociology, medicine, and so forth (e.g., Gould, 1981). Appeals to a "gold standard" of scientific evidence were central to the work of the British Royal Society in the 18th and 19th centuries (Kenner, 1985), with "scientific" evidence subsequently applied to the social problems such as population control and migration, sexuality, birth, marriage and the family, race, and gender relations, poverty and crime, industrial and ecological regulation (Rose, 1999). We find further appeals for the "scientific" management of the state and its institutions across the political spectrum in the 20th century, with heated disputes in the inter- and postwar period between logical positivists and pragmatists, socialists, liberal humanists, and conservatives about what kinds of science should count in the regulation of everyday life. Many contemporary research organizations have their roots in Cold War government funding of the intelligence and research and development sectors (Reich, 2005).

The implications of the scientific, quasi-scientific, and pseudo-scientific control of institutions and everyday life is a prominent theme across millennia of utopian and dystopian fiction from the works of Swift to Orwell and Huxley to recent novels by William Gibson and Margaret Atwood. And we encounter it again in ongoing disputes over the contentious use of "evidence-based policy" and "policy-based evidence" as the rationale for geopolitics, war, and invasion (Haas, 2009) and, most obviously, in current debates over the evidence of human agency in climate change. In a democratic state, making decisions about institutions, communities, environment and place, and people based on rationally argued and supported grounds is a clear imperative—but which scientific evidence, whose sciences and whose interpretations will count is a far tougher call.

The use of evidence and science to address issues of educational equity and social justice is not straightforward. Educational systems have been profoundly troubled by complexity, diversity, and difference. Some of these matters arise in an era of cultural and economic globalization: emergent forms of cultural and ethnic identity and affiliation, large- and medium-scale migration, refugees from warfare and cultural

conflict, employment-driven population shift and movement within countries and regions, and, for many, new and difficult economic conditions. In other instances, it is because of the official and local recognition of the unresolved issues of cultural and linguistic diversity: where communities' voices, rights, and very existence is "named" for recognition where it might have been written over or silenced before. All contemporary educational systems are dealing with increasingly heterogeneous populations. What were previously termed *nonmainstream* and *minority* populations are frequently the norm rather than the exception.

The task at hand also involves addressing the powerfully articulated educational rights of those communities that historically have sat at the margins—often unnamed in curriculum, policy, and practice (see Welner, Chapter 3; Sunderman, Chapter 7; Jordan, Chapter 5). This includes, but is not limited to, cultural and linguistic "minorities" that, in some jurisdictions, have become majorities: that is, Indigenous peoples whose histories, cultures, and languages were excribed from models of curriculum and pedagogy, the specific educational needs and aspirations of girls and women, students of diverse sexual orientations, and those students whose specialized learning needs hitherto have escaped "naming," recognition, and equitable provision.

This is not to invoke the specter of what the media refers to as "political correctness"—nor to sideline the broad goals of just and equitable schooling built around concepts of intergenerational transmission of common and shared knowledge, social skill, and economic capacity, a laudable goal at the heart of democratic education. But it is a practical matter of how school systems can best build on and from the complex forms of what we could broadly term the increasing *epistemological diversity* that teachers, curriculum planners, teacher educators and, indeed, educational policymakers face on a daily basis.

On its surface, evidence-based policy is the new common sense of government secretaries, ministers, advisors, senior civil servants, local school board members and principals. But it calls into question two issues: (a) what might count as "scientific" evidence sufficient to shape rational, normative defensible policy decisions and (b) the interpretive and contextual work of policy formation (see Sunderman, Chapter 7, Wiseman, Chapter 1). Claims that this is a simple, self-evident, or straightforward matter that can be taken without reflection and interpretation, robust dialogue and debate, due consideration of history and theory, narrative as well as expository knowledge, and an understanding the disciplinary bases and assumptions of different educational sciences is naïve and risky. Our authors here provide catalytic reviews and analyses, which, hopefully, will have the effect of unsettling elements of prevailing common sense and provide a broad map of possible research and policy directions.

Alexander Wiseman (Chapter 1) begins the volume by a comparative national and international review of approaches to evidence and equity. He attributes the move toward evidence-based policy to the ongoing demand for political legitimacy, with variable normative goals tending to focus on "quality," "equity," and "control." Wiseman goes on to track how these different goals have been manifest in U.S. states' moves toward evidence-based policy, then contrasting these with other national

and transnational strategies around gender equity. He concludes that the effects of evidence-based policies have tended to reflect specific systems' normative goals, but notes the tendency of systems to adopt a technical bureaucratic approach the definition, collection and interpretation of evidence rather than the "evidence in context" approach advocated by Whitty (2006) and others.

The push toward evidence-based policy has indeed raised major issues around the actual locus of control in educational governance and around the institutional restructuring of state education. Focusing on policy and court cases in the U.S. context, Gail Sunderman (Chapter 7) shows current approaches to equity have evolved over the past century and a half. This evolution is linked to two broader, contradictory movements: first, toward increased federal control over education, and second, toward the promotion of models of privatization and marketization. Sunderman concludes that the byproduct of school reform has been increased centralized bureaucratic control over education. This is predicated on the assumption, she argues, that "standards, assessments and accountability reforms" can address problems that are rooted in persistent social and economic conditions.

As Sam Lucas and Lauren Beresford (Chapter 2) argue, policies and evidence work from a priori definitions of human populations, target groups, and sociological assumptions about what might constitute equitable outcomes. Governance and the modern state works through the establishment of specific grids of specification, classificatory schemes and taxonomies for the definition, categorization, and surveillance of human subjects. These categories are far from "natural" or "transparent." What and whom we include in what may seem a common sense category of "race," "language," or "social class" are elementary problems of social science. Lucas and Beresford review the last half-century of empirical approaches to describing and measuring equity. They then map the foundational assumptions of different categories and definitions and affiliated measurement approaches. Their point is that all classificatory schemes of equity cohorts and approaches to the measurement of their performance align themselves with particular models of society and culture.

Although specific kinds of evidence have been used as rationale and legitimation for policy, the evidentiary bases for legal claims about equality of educational opportunity and equity of outcomes are a neglected area of consideration. Court decisions shape the "bottom line" for debates around public accountability, systemic risk and liability, and issues of social and economic justice. Kevin Welner (Chapter 3) shifts our attention to how we might legally demonstrate equal and unequal provision and outcomes. Reviewing legal and judicial decisions around educational equity, Welner focuses on the variable evidence on equality and inequality in "opportunities to learn." He focuses on the legal precedents in cases where plaintiffs have contested placements and effects of streaming and tracking systems. He concludes that the push for accountability with explicit equity targets has opened the way for "classroom based litigation," which may provide communities and individuals with leverage in defining what might count as equitable learning opportunities and, indeed, outcomes.

Beginning from the *Brown* decision, Will Jordan (Chapter 5) examines whether and how test-based accountability strategies such as *No Child Left Behind* have explained and addressed inequitable patterns of achievement for "people-of-color." Beginning from recognition of the complexity of definition, Jordan questions whether test driven accountability policy has engaged with a rich and diverse research literature on effective practices for minority learners. Jordan reviews the qualitative, ethnographic, and longitudinal pathway research on the schooling of cultural minorities that is often excluded as evidence from policy debates. On this basis, he questions whether test-driven policy has the capacity to enhance authentic learning, to address longstanding issues of social access and mobility facing minority students, and to ameliorate the systemic causes and challenges of economic and cultural exclusion.

As several chapters have indicated—much of the dispute over "evidence" has been based on the strengths and limits of standardized testing. Major U.S., U.K., Canadian, Australian, and New Zealand policymaking has relied principally on an evidence base of standardized test results, despite what we know about the strengths and limits of standardized testing as an assessment technology (e.g., Klenowski, 2009; Moss, Girard, & Hanniford, 2006). What are the alternatives? James Ladwig (Chapter 4) here reviews curriculum research on "social outcomes." He begins by acknowledging a public and educational consensus that traditional academic measures do not address broader goals and outcomes of schooling: ranging from the moral and cultural, to the social and dispositional. He reviews claims on social outcomes from curriculum fields such as multicultural, citizenship, and ecological/environmental education. Ladwig's discussion concludes with a troubling caveat: Although technologies to assess a broader range of educational capacities and performance are available, would we want to bring this host of "other" educational outcomes under the umbrella of high-stakes accountability systems? With what educational consequences and in whose interests?

Dylan Wiliam (Chapter 8) further examines technical issues of equity in assessment. Wiliam begins with an historical review on foundational and operational definitions of "construct" and "validity." Examining three specific cases—higher education admission assessment, the movement toward portfolio-based authentic assessment and assessment-for-learning models, and finally, the impacts of high-stakes testing systems on special needs students—Wiliam makes the case that there has been an incremental conflation of construct and validity. This, in turn, has confounded the interpretation of patterns of gender, cultural/linguistic minority, and special needs achievement, with more attention needed to "construct interpretation" in the analysis of achievement. He argues that conventional and alternative assessment systems need to begin from and rigorously maintain a careful separation of definition of constructs from the test construction process.

The foregoing debates around evidence and equity raise questions about the nature of evidence, the defining and classification of populations, the technical limits of current assessment practices, and the particular political, socioeconomic, cultural, and legal contexts for policy formation and school reform. Yet they focus principally

on equitable access to traditional print-based skills and knowledge. The impact of digital technologies on dominant modes of information, everyday expression and creativity complicate the picture further. The emergence of new technologies challenges the mandate and responsibility of educational systems—raising issues of equitable access and engagement that extend beyond print basics and access to traditional curriculum knowledge. One of the principal contradictions of the current curriculum settlement is that it extols the necessity of new capacities for what are variously termed *information societies* and *knowledge economies* (e.g., creativity, collaboration, critical thinking, intercultural communication), while judging and assessing the efficacy of schools, teachers and instructional approaches at delivering performance via print-based behaviors on standardized tests. Mark Warschauer and Tina Matuchniak (Chapter 6) review the issues around equity raised by the emergence of digital technology. They examine the evidence on home, community, and school technological access and use and they critically reappraise the literature on differential educational outcomes for equity groups. There is, indeed, evidence of a significant and persistent divide in access and achievement, but as importantly, Warschauer and Matuchniak describe an emergent literature on the possibilities of enlisting new technologies to engage marginalized learners and build from their out-of-school engagement with digital culture. They also raise questions about how to assess the overall impacts of new technologies on the achievement of linguistic and cultural minorities, and students from low socioeconomic communities and families.

The first wave of evidence-based policy in the United States and the United Kingdom had a strong concentration on the measurement of student outcomes— often to the exclusion of systematic evidence on teachers' work and career pathways, teacher education, changing workforce demographics, and variable school/community contexts. This is ironic, given the tendency of public debate to define and position teachers both as the problem, and, with an emergent policy focus on "teacher quality," to reposition them as the solution. Judith Warren Little and Laura Bartlett (Chapter 9) review the extensive qualitative and quantitative evidence on teachers' work and its relationship to educational equity. Their argument is that much of the policy literature approaches teacher capacity from an "individualistic frame of reference," which tends to seek "human capital" and "market solutions" to the issue of teacher quality. These obscure the equity implications of the organizational and institutional structures of teachers' work. As a result, the prevailing policy logic tends to stress academic preparedness, workforce diversity, and capacity to engage with cultural diversity, for example, independent of issues of the reform of structural and organizational contexts of schooling. Warren Little and Bartlett conclude by proposing a more comprehensive, contextual model of teacher quality and teaching effectiveness that duly considers the complex institutional ecologies of school organization, teachers' work and community contexts.

A second key omission in current debates over evidence is a focus on schools and communities as social and geographic spaces: teachers and learners do not work in universal, neutral, and generalizable cultural and social environments (Ercikan &

Roth, 2008). Beginning from social geographic research that documents the spatialized character of poverty, social, and cultural marginalization, Kevin Leander, Nathan Phillips, and Katherine Taylor (Chapter 10) shift the lens to another neglected area of policy focus: the social organization of space in learning. Educational administration and policy entails, inter alia, the designing and allocation of space and the assignment, organization, and movement of bodies across that space. Much educational policy, Leander, Phillips, and Taylor argue, has been dominated by a "classroom as container" metaphor—working with the confines of space–time organization established with the founding of the industrial school, and taking for granted the generalizability of the classroom as a unit of description and analysis. Reviewing education research over the past decade that has adopted a social geographic perspective, they raise questions about the impacts of the shaping, structuring, and organization of "schools" and "classrooms" on learning and cultural identity. Like Warschauer and Matuchniak (Chapter 6), they also comment on how minority and lower socioeconomic students are actively using the virtual spaces of new technologies for learning and identity work.

Taken together, these chapters open a broad canvas of what might count as evidence: qualitative and quantitative, empirical and interpretive, synchronic and diachronic, systemic and local. Though none would purport to provide a comprehensive map of the field, each leads us to evidence that has been neglected in current educational debates—including legal, sociodemographic, political economic, sociological, linguistic, anthropological, and social geographic research. The argument here is that the guidance of educational policy and practice committed to equity and social justice requires something more than approaches to accountability reliant on narrow measurement and performance indicators.

Educational policy cannot and does not entail the unmediated, direct translation of factual, empirical claims into direct actions. Even in an idealized moment without overt political pressures and influences—the translation of the empirical, expository claims about "what is" in terms of student outcomes, teacher capacity, school-level operations into the normative, prescriptive moves of legislation, centralized or local intervention in policy and practice is a contingent process. Policy formation requires the building and testing of narrative scenarios (Luke & Woods, 2008). It is contingent on our understandings of how, regionally and locally, something might be seen to be "working," for whom and in whose interests. It is contingent on how one names, identifies, and defines different human subjects. It is contingent on the validity and, indeed, reliability of available evidence. In this context, the diversity of kinds and levels of research that are admitted as "evidence" to the policy bar is a crucial matter. As our contributors here note, the evolution of evidence-based educational policy over the past two decades has been marked by a narrowing, rather than widening of that bar.

At the same time, the translation of "facts to norms" is an interpretive, hermeneutic and, ultimately, speculative and risky process (Habermas, 1996)—no matter how convinced one is of the social facts. There is always a range of plausible interpretations of factual claims and their textual representations, explanations of how they

configure and explain each other. And there is no policy handbook or checklist on how to then translate evidence into complex narrative scenarios that might enable or disenable more equitable, just, and democratic education. These translations, the authors consistently remind us, are utterly dependent on the system's normative goals underlying the use of evidence. And even then they depend up another neglected matter in a policy and media environment that tends to freely assign blame to principals, teachers, students and communities, academics and teacher educators: the actual technical and professional capacities of complex and often unwieldy bureaucracies to lead, resource, and implement reform for equity.

There are other lessons in this volume. To address matters of scalability of reform requires due consideration of the complex mediations involved in the formation and implementation of policy. The bridge between policy and more equitable educational outcomes, broadly construed, are the message systems of curriculum, pedagogy, and assessment. These by definition require detailed, careful consideration for any centrally mandated reform that expects to make a difference. Across this volume, our contributors call for expanded consideration of studies of community cultural and economic contexts, staffrooms and classrooms as mediating "spaces" for more equitable educational practice. By definition, teaching and learning, equity and inequity are situated, historically dynamic, and local phenomena.

Evidence matters in the ongoing struggle for more equitable and just education. But there is no direct link between "fact" and norm, between science and policy. To address questions of equity requires rich, interpretive, and evolving sciences, not a narrow technical approach that invites capture by particular doctrinal and generic approaches to systems reform, public policy, and institutional governance.

ACKNOWLEDGMENTS

NOTES

[1]See AERA's *Standards for Reporting Empirical Social Science Research in AERA Publications* and *Standards for Reporting Humanities-Oriented Research in AERA Publications*, http://www.aera.net/publications/Default.aspx?menu_id=32&id=1850.

REFERENCES

Bryk, A. S., & Gomez, L. (2007, October). *Ruminations on reinventing an R&D capacity for educational improvement*. Paper presented at the American Enterprise Institute Conference: The Supply Side of School Reform and the Future of Educational Entrepreneurship, Washington, DC. Retrieved December 7, 2009, from www.aei.org/event1522

Ercikan, K., & Roth, W. (Eds.). (2008). *Generalizing from educational research*. New York: Routledge.

Gadsden, V., Artiles, A. J., & Davis, J. E. (2009). Risk, equity and schooling. *Review of Research in Education*, *33*, vii-xi.

Green, J., Camilli, G., & Elmore, P. B. (2006). *Handbook of complementary methods in education research*. Mahwah, NJ/Washington, DC: Lawrence Erlbaum/American Educational Research Association.

Gould, S. J. (1981). *The mismeasure of man*. New York: W. W. Norton.

Haas, R. (2009). *War of necessity, war of choice: A memoir of two Iraq wars*. New York: Simon & Schuster.

Habermas, J. (1996). *Between facts and norms* (W. Rehg, Trans.). Cambridge: MIT Press.

Kenner, H. (1985). *The counterfeiters*. Baltimore: Johns Hopkins University Press.

Klenowski, V. (2009). *Raising the stakes: The challenges for teacher assessment.* Paper presented at the annual meetings of the Australian Association for Research in Education, Canberra.

Luke, A., & Woods, A. (2008). Policy and adolescent literacy. In L. Christianbury, R. Bomer, & P. Smagorinsky (Eds.), *Handbook of adolescent literacy research* (pp. 197-215). New York: Guilford Press.

Moss, P., Girard, B. J., & Hanniford, L. C. (2006). Validity in educational assessment. *Review of Research in Education 30*, 109-162.

Parker, L. (2007). Difference, diversity and distinctiveness in education and learning. *Review of Research in Education*, *31*, xi-xv.

Reich, G. A. (2005). *How the Cold War transformed the philosophy of science*. Cambridge, UK: Cambridge University Press.

Rose, N. (1999). *Powers of freedom*. Cambridge, UK: Cambridge University Press.

Whitty, G. (2006). Education(al) research and educational policy making: Is conflict inevitable? *British Educational Research Journal*, *32*, 159-176.