Causal Explanation, Qualitative Research, and Scientific Inquiry in Education

by Joseph A. Maxwell

A National Research Council report, *Scientific Research in Education,* has elicited considerable criticism from the education research community, but this criticism has not focused on a key assumption of the report—its Humean, regularity conception of causality. It is argued that this conception, which also underlies other arguments for "scientifically-based research," is narrow and philosophically outdated, and leads to a misrepresentation of the nature and value of qualitative research for causal explanation. An alternative, realist approach to causality is presented that supports the scientific legitimacy of using qualitative research for causal investigation, reframes the arguments for experimental methods in educational research, and can support a more productive collaboration between qualitative and quantitative researchers.

A major effort to establish "scientifically-based research" (SBR) in education has been under way for some time (Feuer, Towne, & Shavelson, 2002). Key manifestations of this effort are found in the No Child Left Behind Act and the reauthorization and reorganization of the Office of Educational Research and Improvement as the Institute for Educational Science (Eisenhart & Towne, 2003). The most detailed presentation of what SBR would mean for education research is the report by the National Research Council (NRC), *Scientific Research in Education* (NRC [2002] report). This work, unlike some others that use "scientifically-based research" mainly as a code word for randomized experiments, calls for a broad definition of science that encompasses a range of approaches and methods, both qualitative and quantitative, which is consistent with the American Educational Research Association's (2003) "Resolution on the Essential Elements of Scientifically-based Research."

Despite this, the NRC (2002) report has provoked substantial criticism and concern among education researchers (Eisenhart & Towne, 2003). This concern is evidenced by the recent publication in the *Educational Researcher* of several critical commentaries on the report (Berliner, 2002; Erickson & Gutierrez, 2002; St. Pierre, 2002) as well as by a number of critical addresses and sessions devoted to the report at the 2003 AERA meeting (e.g., Lather, 2004).

However, none of the criticism of the NRC (2002) report, or of SBR generally, has systematically challenged a key feature of SBR, the claim that experimental methods are the preferred strategy, the "gold standard," for causal investigations.[1] In this article, I argue that this claim has been based on a restrictive and philosophically problematic model of causality, a model that is implicit in most of the advocacy for SBR. I focus on the NRC report because it (a) is the most detailed and sophisticated presentation of SBR and the assumptions underlying this approach, (b) has been the subject of considerable discussion in the educational research community, and (c) is clearly intended to justify a broadly based definition of science. My critique applies even more forcefully to less sophisticated or less ecumenical proposals for SBR in education.

The view of causation that I am challenging, often referred to as the "regularity" approach, holds that we can't directly observe causation, only the regularities in the relationships between events. I present an alternative view, associated with the position in the philosophy of science often termed "scientific realism," that involves a fundamentally different model of causal explanation.[2] This model challenges the privileged position that SBR gives to randomized experiments in education research, and the concomitant dismissal of qualitative research as a rigorous means of investigating causality. I argue that a realist understanding of causality is compatible with the key characteristics of qualitative research, and supports a view of qualitative research as a legitimately scientific approach to causal explanation. This view can promote a broader use of qualitative research in education research, and a more productive collaboration between qualitative and quantitative researchers.

The NRC (2002) report's argument for the pre-eminence of randomized experiments in causal investigation involves a series of linked assumptions and claims about causality and the appropriate methods for investigating this. Specifically, the NRC report:

1. Assumes a regularity view of causation.
2. Privileges a variable-oriented approach to research over a process-oriented approach.
3. Denies the possibility of observing causality in single cases.
4. Neglects the role of context as an essential component of causal explanation.
5. Neglects the importance of meaning for causal explanation in social science.
6. Asserts that qualitative and quantitative research share the same logic of inference.
7. Presents a hierarchical ordering of methods for investigating causality, giving priority to experimental and other quantitative methods.

I discuss each of these assumptions in turn, showing how they are manifested in the report, and present a realist alternative to

Educational Researcher, Vol. 33, No. 2, pp. 3–11

these assumptions and the implications of this alternative for education research.

The Regularity View of Causation

The regularity conception of causality derives from David Hume's analysis of causality,[3] as elaborated by John Stuart Mill and further developed by positivist philosophers such as Carl Hempel and Ernest Nagel. Hume's fundamental idea was that we cannot directly perceive causal relationships, only what he called the "constant conjunction" of events. He denied that we can have any knowledge of causality beyond the observed regularities in associations of events, and proscribed any reference to unobservable entities and mechanisms. While the restriction to observable entities was abandoned with the collapse of logical positivism, the idea that causality is fundamentally a matter of regularities in our data was the standard view in philosophy of science for much of the 20th century (Salmon, 1989; see also House, 1991). This view was codified by Hempel and Oppenheim (1948) in what they called the "deductive-nomological" theory of causation; Salmon (1989) labels the current manifestation of this approach the "unification" view. This approach "seeks laws and principles of the utmost generality" (Salmon, 1989, p. 182), and sees scientific explanation as the fitting of specific facts into this framework of laws; particular events are explained in terms of these laws, which may be statistical as well as deterministic. This view treats the actual process of causality as unobservable, a "black box," and focuses on discovering whether there is a systematic relationship between inputs and outputs. This conception of causality is "the basis of ordinary quantitative research and of the stricture that we need comparison in order to establish causality" (Mohr, 1996, p. 99).

The central manifestation of the regularity view in the NRC (2002) report is its presentation of causality as primarily pertaining to *whether* x caused y, rather than *how* it did so (p. 110). The NRC report presents a sequence of three types of research questions: (a) description—What is happening? (b) cause—Is there a systematic effect?, and (c) process or mechanism—Why or how is it happening? (p. 99). Although the report devotes some space to the investigation of mechanisms (pp. 117–123), it uses "cause" mainly for the systematic relationship between variables, rather than for causal processes. "Knowledge about causal relationships" is equated with "estimating the effects of programs" (p. 125), and the phrases "causal questions" and "causal studies" primarily refer to the investigation of causal effects rather than the search for causal mechanisms.

The most widely accepted alternative to the regularity approach to causality is a realist approach that sees causality as fundamentally referring to the actual causal mechanisms and processes that are involved in particular events and situations. For the philosophy of science in general, this approach has been most systematically developed by Salmon (1984, 1989, 1998), who refers to it as the "causal/mechanical" view. This approach

> makes explanatory knowledge into knowledge of the . . . mechanisms by which nature works . . . . It exhibits the ways in which the things we want to explain come about. This way of understanding the world differs fundamentally from that achieved by way of the unification approach. Whereas the unification approach is 'top-down', the causal/mechanical is 'bottom-up.' (Salmon, 1989, pp. 182–183)

Salmon argues that the two approaches "may be taken as representing two different, but compatible, aspects of scientific explanation" (1989, p. 183). However, he considers only the causal/mechanical view to be *causal* explanation (1998, p. 9).

For the social sciences, this approach to causation has been developed most extensively (but by no means exclusively) by those calling themselves "critical realists" (e.g., Bhaskar, 1978, 1989; Harre, 1975; cf. Archer, Bhaskar, Collier, Lawson, & Norrie 1998). Sayer (1992) argues that

> much that has been written on methods of explanation assumes that causation is a matter of regularities in relationships between events, and that without models of regularities we are left with allegedly inferior, 'ad hoc' narratives. But social science has been singularly unsuccessful in discovering law-like regularities. One of the main achievements of recent realist philosophy has been to show that this is an inevitable consequence of an erroneous view of causation. Realism replaces the regularity model with one in which objects and social relations have causal powers which may or may not produce regularities, and which can be explained independently of them. In view of this, less weight is put on quantitative methods for discovering and assessing regularities and more on methods of establishing the qualitative nature of social objects and relations on which causal mechanisms depend. (pp. 2–3)

Thus, despite Salmon's claim that the "unification" and "causal-mechanical" approaches are equally legitimate as general models for scientific explanation, the "law-like regularities" assumed by the unification approach (including statistical laws analogous to those of quantum mechanics) clearly have not been established in the social sciences (Bohman, 1991, pp. 18–30; Sayer, 1992, pp. 125–129). For the social sciences, the unification approach lacks the prerequisite knowledge to provide a powerful explanatory tool.[4] This suggests that realist, process-oriented investigations deserve a more prominent place in educational research, including experimental research. Shadish, Cook, and Campbell (2002), in what is arguably the most detailed and sophisticated presentation of the case for experimental research, state that

> the unique strength of experimentation is in describing the consequences attributable to deliberately varying a treatment. We call this *causal description*. In contrast, experiments do less well in clarifying the mechanisms through which and the conditions under which that causal relationship holds—what we call *causal explanation*. (p. 9)

Referring to a "delicate balance" between causal descriptions and causal explanations, they assert that "most experiments can be designed to provide better explanations than is the case today" (p. 12), and describe several studies in which qualitative methods were used to substantially strengthen the understanding of causal mechanisms in experimental investigations (pp. 390–392).

A Variable-Oriented Approach to Research

The regularity assumptions in the NRC (2002) report are characteristic of what Mohr (1982) has called "variance theory," an approach that he distinguishes from "process theory." Variance theory deals with variables and the correlations among them; it is based on an analysis of the contribution of differences in values of particular variables to differences in other variables. The comparison of conditions or groups in which the presumed causal factor takes different values, while other factors are held constant or statistically controlled, is central to this approach to causation, and is affirmed as "a fundamental scientific concept in making causal claims" by the NRC report (p. 110). Thus, variance theory tends to be associated with research that employs experimental or correlational designs, quantitative measurement, and statistical analysis. As Mohr (1982) notes, "the variance-theory model of explanation in social science has a close affinity to statistics. The archetypal rendering of this idea of causality is the linear or non-linear regression model" (p. 42).

Process theory, in contrast, deals with *events* and the processes that connect them; it is based on an analysis of the causal *processes* by which some events influence others. It is fundamentally different from variance theory as a way of thinking about scientific explanation. Pawson and Tilley (1997), in their realist approach to program evaluation, state that

> When realists say that the constant conjunction view of one event producing another is inadequate, they are not attempting to bring further "intervening" variables into the picture . . . . The idea is that the mechanism is responsible for the relationship itself. A mechanism is . . . not a variable but an *account* of the makeup, behaviour and interrelationship of those processes which are responsible for the regularity. (pp. 67–68)

Similar distinctions have been presented by many other writers, and appear to have been independently invented in a number of disciplines. These include the distinctions between variable-oriented and case-oriented approaches (Ragin, 1987), propositional knowledge and case knowledge (Shulman, 1990), and factor theories and explanatory theories (Yin, 1993, pp. 15–21). Sayer (1992, pp. 241–251) similarly distinguishes between extensive and intensive research designs; extensive designs address regularities, common patterns, and distributions of features of populations, while intensive designs focus on how processes work in particular cases.[5]

Such distinctions are found in the natural sciences as well. Gould (1989) describes two styles of science, one characteristic of physics and chemistry, the other of disciplines such as evolutionary biology, geology, and paleontology, which deal with unique situations and historical sequences. He argues that

> The stereotype of the "scientific method" has no place for irreducible history. Nature's laws are defined by their invariance in space and time. The techniques of controlled experiment, and reduction of natural complexity to a minimal set of general causes, presuppose that all times can be treated alike . . . .
>
> But the restricted techniques of the "scientific method" cannot get to the heart of [a] singular event involving creatures long dead on an earth with climates and continental positions markedly different from today's. The resolution of history must be rooted in the reconstruction of past events themselves—in their own terms—based on narrative evidence of their own unique phenomena.
>
> Historical science is not worse, more restricted, or less capable of achieving firm conclusions because experiment, prediction, and subsumption under invariant laws of nature do not represent its usual working methods. The sciences of history use a different mode of explanation, rooted in the comparative and observational richness of our data. (pp. 277–279)

What Gould (1989) calls "historical explanation" is better seen as process explanation, since many applications of this are not strictly historical. This approach is also widely used in ecology, psychology, anthropology, and sociology.

Explicit use of process theory as an alternative to variance theory is becoming more common in the social sciences. For example, Rogers (1995), in what is widely viewed as the pre-eminent work on the diffusion of innovations, argues that the use of variance-theory methods, such as surveys, in diffusion research "is intellectually destructive of the 'process' aspects of the diffusion of innovations" (p. 122). He asserts that

> in order to explore the nature of the innovation-decision process, one needs a dynamic perspective to explain the causes and sequences of a series of events over time. Data-gathering methods for process research are less structured and the data are typically more qualitative in nature. Seldom are statistical methods used . . . . The general point here is that research on a topic like the innovation-decision process must be quite different from the variance research that has predominated in the diffusion field in the past." (p. 189)

Becker (1992, pp. 205–207) argues that many of the problems with variance theory are well known, but tend to be ignored because there is no simple way to deal with them within a variance framework. Education research would benefit from a more explicit recognition and discussion of these problems (e.g., Shadish et al., 2002, pp. 456–504; Yin, 1993, pp. 15–18, 38–39), and the use of process approaches to deal with these.

Observing Causality in Single Cases

The NRC (2002) report is committed, in two ways, to viewing causal explanations as necessarily general. First, the regularity model of causation implies that causality can never be identified in single events or cases, only through repeated observations of a relationship between two variables or events. Second, the report consistently emphasizes that causal explanations are intrinsically about "systematic effects" rather than single events (p. 108), and that the goal of scientific research is to "replicate and generalize across studies" (p. 52). Scientific theories are described as "conceptual models that explain some phenomenon" (p. 59), but the authors then state that "some research in the social sciences seeks to achieve a deep understanding of particular events or circumstances rather than theoretical understanding that will generalize across situations or events" (p. 59). In the report, in-depth studies of particular settings are consistently relegated to descriptive rather than causal or process research (e.g., pp. 105–108, 120–122). The authors state that such studies can generate plausible hypotheses about causal relationships, and can be valuable "when used in conjunction with causal methods" (p. 106), but "must advance to other levels" (p. 108) to answer questions about "what works."

A realist, process-oriented conception of causal explanation entails a quite different approach to understanding particular events or situations. Salmon (1998) states that

> Hume concludes that it is only by repeatedly observing associated events that we can establish the existence of causal relations. If, in addition to the separate events, a causal connection were observable, it would suffice to observe one case in which the cause, the effect, and the causal relation were present. (p. 15)

Salmon (1998) then argues that "causal processes are precisely the connections Hume sought, that is, that the relation between a cause and an effect is a physical connection" (p. 16).[6] Putnam (1999, pp. 140–141) also claims that we can observe causation, quoting Anscombe:

> First, as to the statement that we can never observe causality in the individual case. Someone who says this is just not going to count anything as "observation of causality." This often happens in philosophy; it is argued that "all we find" is such-and-such, and it turns out that the arguer has excluded from his idea of "finding" the sorts of things he says we don't "find." (Anscombe, 1971, p. 137, quoted in Putnam, 1999, p. 141)

Ducasse (1926) and Davidson (1967) have also argued that causation can be identified in a single case, and Cartwright (2000) claims that regularity approaches cannot provide adequate accounts of causation without presupposing singular causal knowledge.

Shadish et al. (2002), despite their tendency to rely on variance theory in their explication of experimental methods, accept Davidson's and others' arguments for the identification of causes in single instances (p. 465). They do so because they believe that experimenters "*can* infer cause in single experiments . . . [and that] providing them with conceptual help in doing so is a virtue, not a vice; failing to do so is a major flaw in a theory of cause-probing methods" (p. 465).

A long line of experimental research on perception supports the view that we can, in fact, observe causal relationships. Sperber (1995), in his introduction to a book on causal cognition, states that the "anti-Humean idea, that causal relations may be perceived and not just inferred, has been espoused, developed, and experimentally tested by the Belgian psychologist Andre Michotte (1946), and is elaborated on in several contributions to this book" (p. xvi). In fact, experimental research supports a distinction (in animals as well as humans) between what has been called "natural" causal perception in a single event, and "arbitrary" causal judgment based on the identification of regularities (Dickinson & Shanks, 1995; Kummer, 1995). These two types of causal knowledge, which closely correspond to the distinction presented above between realist and regularity conceptions of causation, appear to involve quite different cognitive processes (Premack & Premack, 1995). While causal perceptions (like all perceptions and judgments) may be mistaken, they are not *necessarily* mistaken, and this concept, qualified by the idea of "epistemically ideal conditions" (Putnam, 1990, p. vii), further undermines the Humean argument against the observation of causation in particular instances.

Miles and Huberman (1994), drawing on their experience with multisite education policy research, use the phrase "local causality" to express the practical application of this idea of particular causation. They provide a detailed explanation of how qualitative methods can provide credible accounts of such causal relationships and processes (pp. 144–165; see also Maxwell, in press; Weiss, 1994, pp. 179–181), arguing that "qualitative analysis, with its close-up look, can identify *mechanisms*, going beyond sheer association. It is unrelentingly *local*, and deals well with the *complex* network of events and processes in a situation" (p. 147). And they conclude that "if you've done it right, you will have respected the complexity of local causality as it has played out over time, and successfully combined 'process' and 'variable' analysis" (p. 160).

This argument dovetails well with the NRC (2002) report's "Scientific Principle 3: Use methods that permit direct investigation of the question" (p. 3). Experimental researchers relying exclusively on a regularity model of causation assume, following Hume, that the researcher can't directly observe causation, and therefore must depend on inferring causal relationships from measured covariation of variables. Qualitative studies that are based on a process approach to causation, in contrast, attempt to directly investigate causal mechanisms. This argument has been made not only by self-conscious realists, but also by researchers with no explicit commitment to this position. For example, Britan (1978) claims that "experimental evaluations relate program treatments to program effects without directly examining causal processes, [while] contextual [qualitative] evaluations investigate causal relationships . . . by directly examining the processes through which results are achieved" (p. 231). And Jankowski (1991), reporting on his 10-year participant observation study of urban gangs in three cities, states that

> Unable to observe gang violence directly, researchers have treated it as a dependent variable, something to be explained using structural and individual-oriented independent variables. This study . . . seeks to understand the anatomy of violence as well as to explain it. (p. 138)

As Becker (1996) notes, "It is invariably epistemologically dangerous to guess at what could be observed directly" (p. 58). The ability of qualitative methods to directly investigate causal processes is a major contribution that this approach can make to scientific inquiry in education. Unfortunately, this ability has not only been denied by most quantitative researchers, but also by many qualitative researchers, and specific methods for identifying and verifying causal processes need further explication and development (Maxwell, in press).

Neglect of the Role of Context in Causal Explanation

Realist social researchers place considerable emphasis on the context dependence of causal explanation (e.g., Sayer, 1992, pp. 60–61; Huberman & Miles, 1985, p. 354). Pawson and Tilley (1997) sum up this position in their formula "mechanism + context = outcome" (p. xv). They maintain that "the relationship between causal mechanisms and their effects is not fixed, but contingent" (p. 69); it depends on the context within which the mechanism operates. This is not simply a claim that causal relationships vary across contexts; it is a more fundamental claim, that the context within which a causal process occurs is, to a greater or lesser extent, intrinsically involved in that process, and often cannot be "controlled for" in a variance-theory sense without misrepresenting the causal mechanism (Sayer, 2000, pp. 114–118). Thus, Goldenberg (1992), in a case study of the reading progress of two students and the effects of their teacher's behavior on this progress, states, "If we see these dimensions as variables divorced from this context, we risk distorting the role they actually play" (p. 540). For the social sciences, the social and cultural contexts of the phenomenon studied are crucial for understanding the operation of causal mechanisms.

The NRC (2002) report also argues for the importance of context (pp. 5, 22, 84–91), but primarily for two other reasons. First, the authors emphasize that context is important for the *generalization* of causal statements; "attention to context is especially critical for understanding the extent to which theories and findings may generalize to other times, places, and populations" (p. 5; cf. p. 91). This view treats context as one source of "noise" that "make[s] replication—the key to boosting certainty in results and refining theory—more difficult and nuanced" (p. 83) in social and behavioral research. Second, the report emphasizes the effect of the social and political contexts of education on the conduct of the research itself, as opposed to their causal involvement in the phenomena studied (pp. 80, 85, 91). The authors acknowledge the relevance of context for understanding the influence of student and classroom diversity on learning (p. 90), but this role of context receives little explicit attention in their discussion of causal explanation.

Criticism of experimental research in general, and of the NRC report in particular, for its neglect of context is common (e.g., Berliner, 2002). However, this criticism usually has not distinguished between the importance of context for generalizability (the difficulty of finding regularities across contexts in education research) or for the nature of education research itself, and the fundamental role of context in causal explanations of educational phenomena. To develop adequate explanations of educational phenomena, and to understand the operation of educational interventions, we need to use methods that can investigate the involvement of particular contexts in the processes that generate these phenomena and outcomes.

Neglect of the Importance of Meaning for Causal Explanation

Similarly to context, realist social scientists see the meanings, beliefs, values, and intentions held by participants in a study as essential parts of the causal mechanisms operating in that setting (e.g., Huberman & Miles, 1985, p. 354; House, 1991, p. 6). Sayer (1992) states, "Social phenomena are concept-dependent . . . what the practices, institutions, rules, roles, or relations *are* depends on what they mean in society to its members" (p. 30). Thus, causes are not restricted to physical entities or processes; they can include mental phenomena (Sayer, 1992, pp. 110–112). While there has been much philosophical debate over the legitimacy of seeing mental phenomena as causes of behavior (cf. Heil & Mele, 1993), the work of Davidson[6] (1980, 1993), McGinn (1991), and Putnam (1999) has given this position considerable philosophical credibility. This view of intentions, beliefs, and meanings as causes is fundamental to our common-sense explanations of actions, as well as to psychology as a science (Davidson, 1997, p. 111), and has been affirmed not only by critical realists but by many other philosophers and social scientists as well (e.g., MacIntyre, 1967; Menzel, 1978).

The NRC (2002) report discusses human beliefs, values, and volition mainly in terms of their effect on the education research enterprise, rather than as an integral part of the phenomena studied (pp. 84–87). The report acknowledges that "research on human action must take into account individuals' understandings, beliefs, and values as well as their observable behavior" (p. 16), but this idea receives little subsequent discussion, and its implications for causal investigation are never addressed. The NRC report, and SBR generally, largely ignore the social science traditions that focus on meaning rather than behavior, on what is often referred to as the "interpretive" dimension of social life (Shulman, 1990, pp. 74–76), and, thus, they fail to incorporate the interpretive character of our understanding of human thought and action in their conception of causal explanation (Sayer, 2000, pp. 45–46). Meanings, beliefs, and volitional actions constitute processes that can't be converted to variables, even "intervening" variables, without fundamentally concealing and misrepresenting the nature of this process (Blumer, 1956). The investigation of the interpretive dimension of social phenomena normally requires qualitative methods, which are particularly well suited for elucidating participants' actual perspectives and interpretations. Ignoring this dimension can lead to serious distortions of causal conclusions.

The Assertion that Qualitative and Quantitative Research Share the Same Logic of Inference

The NRC (2002) report asserts that "all rigorous research—quantitative or qualitative—embodies the same underlying logic of inference" (p. 67), and the authors state that they believe the distinction between these two approaches is "outmoded" (p. 19). The source cited for the first of these claims is King, Keohane, and Verba (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*. However, this work is based almost entirely on variance theory, and bears little resemblance to qualitative research as this term is normally understood. Mohr (1996), in a generally positive discussion of the book and with no apparent ironic intent, describes it as

> an excellent book on the constructive use of quantitative reasoning in small-n research . . . For the most part, these [procedures] involve considering X and Y to be *variables*, maintaining an orientation toward factual [regularity] causal reasoning. (pp. 120–121)

Thus, King et al. (1994) state that "precisely defined statistical methods that undergird quantitative research represent abstract formal models applicable to all kinds of research, even that for which the variables cannot be measured quantitatively" (p. 6). Since their goal is to get qualitative researchers to use quantitative, variance-theory strategies, it is not surprising that they argue that the differences between the qualitative and quantitative traditions are "only stylistic and . . . methodologically and substantively unimportant" (p. 4). This position would be rejected by almost all qualitative researchers.

The NRC (2002) report also claims that qualitative and quantitative research are epistemologically similar (p. 19), citing Howe and Eisenhart (1990) as well as King et al. (1994). However, Howe and Eisenhart's argument is that framing the difference between qualitative and quantitative research in terms of two "paradigms," positivism and interpretivism, is outdated, since positivism is no longer a credible philosophical position. They argue that all research must be justified within a nonpositivist framework, and deny that the "vestigial methodological positivism" often associated with quantitative research can be epistemologically justified (p. 6). In addition, they repeatedly emphasize that different education research traditions have different "logics in use" and, thus, different specific standards (pp. 3–4, 8); "there can, accordingly, be no monolithic unity of scientific method"

(p. 3). Thus, Howe and Eisenhart's arguments provide no support for the NRC report's claim that there are no important differences in logic between qualitative and quantitative research (see also Becker, 1996; Bohman, 1991, p. 39; Sleeter, 2001). Understanding these differences in logic, and their implications for the relative strengths and limitations of the two approaches, is essential for productively utilizing both approaches.

A Hierarchical Ordering of Methods for Assessing Causality

This assumption of the fundamental similarity between qualitative and quantitative methods, when combined with the regularity and variance-theory assumptions of the NRC (2002) report, leads the authors to impose standards that are appropriate for quantitative or experimental research on *all* research, in the name of "scientific rigor." For example, the authors recognize that qualitative research typically proceeds in a relatively inductive manner, developing theory and testable interpretations during the study rather than exclusively in advance (p. 62, fn. 4). However, they assert that "more rigorous studies will begin with more precise statements of the underlying theory . . . and will generally have a well-specified hypothesis before the data collection and testing phase is begun" (p. 101). Studies that lack these "may still be scientific, although they are obviously at a more rudimentary level" (p. 101).

These assumptions lead to serious misunderstandings of qualitative research and to a hierarchical prioritizing of quantitative and experimental methods for explanatory purposes. The latter is conveyed both explicitly, through the ranking of their three types of research questions, and implicitly, through the kinds of examples chosen to illustrate "strong" and "weak" designs and the language used to describe these. Almost all of the qualitative research mentioned in the report is treated as "descriptive," addressing the first question, "What is happening?" While such studies "have clearly contributed important insights to the base of scientific knowledge" (NRC, 2002, p. 108), they are clearly identified as a lower category by phrases such as "advance beyond the descriptive phase toward more precise scientific investigation of causal effects and mechanisms" (p. 101).

In contrast, all of the examples of research addressing the second question, "Is there a systematic effect?" are of experimental, quasi-experimental, or quantitative model-fitting studies (NRC, 2002, pp. 110–117), with randomized experiments described as "the best way to answer such questions" (p. 102). Although the use of multiple qualitative methods is described as helpful in ruling out alternative interpretations and in generating plausible causal hypotheses, these methods are recommended only in conjunction with quantitative or experimental designs (p. 109).

The third question, "Why or how is it happening?" is the one for which process-oriented qualitative research could be most valuable. However, because qualitative research has been relegated to the "descriptive" category, such studies are described as appropriate only when theory is weak. Because they have a "strong descriptive component," they can "illuminate unforeseen relationships and generate new insights" (NRC, 2002, p. 120), but even "scientifically rigorous case studies . . . cannot (unless combined with other methods) provide estimates of the likelihood . . . that they have identified the right underlying cause" (p. 108). All of the studies of mechanism using qualitative methods that are discussed in the report (pp. 117–123) also employed variance-oriented strategies, and it is the latter strategies (e.g., statistically holding constant other variables) that are claimed to "explain the mechanism" (p. 119).

In contrast, Abbott (1992) and Becker (1992) provide detailed accounts of how a reliance on variance theory distorts sociologists' causal analyses of cases, and they argue for a more systematic and rigorous use of narrative and process analysis for causal explanation. Abbott (1992) asserts that

> a social science expressed in terms of typical stories would provide far better access for policy intervention than the present social science of variables . . . . Anybody who knew the typical stories of organizations under great external and internal stress would never have believed that breaking up AT&T would result in a highly profitable firm and a cheaper overall phone service. But policymakers saw economists' equations proving that profit equaled so many parts research plus so many parts resources plus so many parts market competition and so on. No one bothered to ask whether one could tell a real story that led from AT&T as of 1983 to the vision they had in mind. (pp. 79–80)

The misunderstanding and subordination of qualitative research documented above is harmful to education research in two ways. First, it inhibits researchers from fully using the unique advantages of qualitative methods to strengthen their causal investigations. Second, it damages the relationships between qualitative and quantitative researchers, making genuine collaboration more difficult.

Conclusions

To summarize, I have argued that the NRC (2002) report (and SBR generally) assumes a regularity, variance-oriented understanding of causation, and ignores an alternative, realist understanding. This leads the authors to ignore or deny the possibility of identifying causality in particular cases, the importance of context as integral to causal processes, and the role of meaning and interpretive understanding in causal explanation—all issues for which qualitative research offers particular strengths. In addition, the NRC report denies that there are important differences in epistemology and logic between qualitative and quantitative research. In combination, these assumptions lead to a hierarchical ordering of research methods in the report, treating qualitative methods as merely descriptive and supplementary to "causal," quantitative methods, largely ignoring the unique contributions that qualitative methods can make to causal investigation. I believe that this is one important reason why there has been such a negative reaction to the report, and to SBR more generally, by many educational researchers.

I also argue that the realist alternative that I have presented to the dominant regularity model of causality can provide a way out of the somewhat polarized confrontation between qualitative and quantitative researchers on this issue of causal investigation. A realist, process-oriented approach to explanation is compatible with, and facilitates, the key strengths of qualitative research. In particular, it recognizes the reality and importance of *meaning*, as well as of physical and behavioral phenomena, as having explanatory significance, and the essentially *interpretive* nature of our understanding of the former. It also recognizes the explanatory importance of the *context* of the phenomena studied, and does so in a way that does not simply reduce this context to a set of "extraneous variables." It relies fundamentally on an understanding of the *processes* by which an event or situation occurs, rather than simply a comparison of situations involving the presence and absence of the presumed cause. Finally, it legitimates a concern with understanding *particular* situations and events, rather than addressing only general patterns.

However, in contrast to most relativist, constructivist, or postmodern positions, which have often been invoked as stances for qualitative research, realism emphasizes the fundamental importance of validity issues (Hammersley, 1992; Maxwell, 1992, 2002, 2004), as well as the legitimacy of causal explanation, in qualitative research. Thus, realism supports the argument that qualitative research can be scientific in the full sense of the term, providing explicitly developed, testable explanations for the phenomena studied.

None of this should be taken as denying or disparaging the value of experimental research in education, or even of the importance of what Shadish et al. (2002) call "causal description." While realists have tended to be critical of experimental research (e.g., Pawson & Tilley, 1997), I believe that realism is a productive stance for both experimental and qualitative research. Donald Campbell, the pre-eminent figure in the experimental tradition in social science, was a realist (Campbell, 1988; cf. Maxwell, 1990), and one of the major works in this tradition (Cook & Campbell, 1979) is explicitly grounded in a realist epistemology. Cook (2002) states that the regularity approach is a "less comprehensive and less esteemed theory of causation" than explanatory approaches (p. 180), and argues, "experiments should be designed *to explain* the consequences of interventions and not just to describe them" (p. 189). He has summed up this argument in the dictum "No more black-box experiments" (p. 181).

Thus, I would reframe the argument for experimental methods in terms of a realist understanding of causation. Rather than randomized experiments being "an ideal method when entities being examined can be randomly assigned to groups" (NRC, 2002, p. 109), let alone the "gold standard" for causal investigation, as some have claimed, I would argue that strictly experimental designs, with no qualitative components, are a comparatively powerful method for understanding causality only when three conditions obtain. First, there should be a well-developed theory that informs the intervention and research design and allows interpretation of the experimental results (Bernard, 2000, pp. 55–56); this is one of the NRC report's conditions for using quantitative methods for exploring mechanism (pp. 118–120). Second, the causal process investigated should be manipulable, fairly straightforward and simple (NRC, 2002, pp. 109, 125; Cook, 2002, p. 179), and relatively free from temporal and contextual variability, a condition that Sayer (1992, pp. 121–125) calls a "closed system." Third, the situation should not be conducive to the direct investigation of causal processes.

These three conditions are often met in medicine and agriculture, fields in which the randomized experiment is a dominant strategy (although in both fields, laboratory investigation of mechanisms is an essential complement to the randomized assessment of the effect of treatments). These conditions are less often met in education, where causal processes are often complex, temporally and contextually variable, and directly observable. In such situations, qualitative methods have distinct advantages for identifying the influence of contextual factors that can't be statistically or experimentally controlled, for understanding the unique processes at work in specific situations, and for elucidating the role of participants' beliefs and values in shaping outcomes.

In many instances, the optimal approach will be to combine qualitative and experimental methods, as both the NRC (2002) report and Shadish et al. (2002) advocate. Shulman (1990) claims that

> when investigators have learned to speak each others' languages, to comprehend the terms in which other programs' research questions are couched, then processes of deliberation over findings can yield the hybrid understandings not possible when members of individual research programs dwell in intellectual ghettos of their own construction. (p. 84)

For this to happen productively, practitioners of both approaches will need to develop a better understanding of the logic and practice of the other's approach, and a greater respect for the value of the other perspective. Thus, my argument is directed to qualitative researchers who deny the legitimacy or relevance of experimentation or causality for education research (cf. Maxwell, 2004) as well as to quantitative researchers who dismiss or distort the potential contributions of qualitative research to the task of developing causal explanations. This mutual understanding is a prerequisite for the optimal integration of qualitative and quantitative approaches in educational research (Johnston & Swift, 1994, p. 71; cf. Maxwell & Loomis, 2003), and can lead to a collaboration between quantitative and qualitative researchers that is based on equality and complementarity rather than hierarchy and epistemological or political dominance by one perspective.

NOTES

I want to thank the editors of *ER*, Evelyn Jacob and Stephen White, for encouraging me to write this article, and the editors, Eamonn Kelly, and four anonymous reviewers for their valuable feedback.

[1] In their response to the NRC (2002) report, Erickson and Gutierrez (2002) devote one paragraph to problems with the Humean conception of causality, clearly identifying some of the difficulties that this conception creates for education research. However, they neither document the role that this conception of causality plays in the argument of the NRC report, nor develop the specific implications of this and alternative conceptions, both for causal inquiry and for the relationships between qualitative and quantitative research methods and approaches.

[2] House (1991) has provided a detailed and perceptive explanation of scientific realism and its relevance to education research, one that anticipates many of my criticisms of the NRC (2002) report.

[3] Not all philosophers believe that Hume, in fact, held a regularity theory of causation. For a detailed argument that Hume was actually a realist with respect to causation, see Strawson (1989).

[4] For a more general critique of "unification" approaches, in both the physical and social sciences, see Cartwright (1999).

[5] Other social scientists (e.g., Becker, 1998, pp. 57–66; Bruner, 1986) make a similar distinction between "logical" or "causal" analysis and "stories" or "narratives." This distinction has the virtue of highlighting the narrative character of process explanations but is much broader than the one I'm presenting here and neglects the ways in which narratives can constitute causal explanations without invoking variables (Abbott, 1992; Becker, 1992).

[6] Note that Salmon is talking about physical science (cf. 1989, p. 180). As discussed below, realists dealing with the social sciences typically consider mental as well as physical processes to be causal, although mental processes are not amenable to direct observation.

[7] Although Davidson is the philosopher usually cited for the argument that reasons are causes, both he and many of his critics accept the view that causal explanation requires causal laws, although Davidson (1975) sees these laws as individual rather than general. The view of reasons as causes becomes less problematic when this dependence of causality on laws is abandoned (cf. Sayer, 1992, pp. 126–127; Putnam, 1999, pp. 73–91, p. 200, fn. 11, 14).

REFERENCES

Abbott, A. (1992). What do cases do? Some notes on activity in sociological analysis. In C. C. Ragin & H. S. Becker (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 53–82). Cambridge, England: Cambridge University Press.

American Educational Research Association (2003). Resolution on the Essential Elements of Scientifically-based Research. Retrieved February 14, 2004, from http://www.aera.net/meeting/councilresolution03.htm

Anscombe, G. E. M. (1971). *Causality and determination: An inaugural lecture*. London: Cambridge University Press. Archer, M., Bhaskar, R., Collier, A., Lawson, T., & Norrie, A. (1998). *Critical Realism: Essential Readings*. London: Routledge.

Becker, H. S. (1992). Cases, causes, conjectures, stories, and imagery. In C. C. Ragin & H. S. Becker (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 53–71). Chicago: University of Chicago Press.

Becker, H. S. (1996). The epistemology of qualitative research. In R. Jessor, A. Colby, & R. A. Shweder (Eds.), *Ethnography and human development: Context and meaning in social inquiry* (pp. 205–216). Cambridge, England: Cambridge University Press.

Becker, H. S. (1998). *Tricks of the trade: How to think about your research while you're doing it*. Chicago: University of Chicago Press.

Berliner, D. (2002). Educational research: The hardest science of all. *Educational Researcher 31*(8), 18–20.

Bernard, H. R. (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Bhaskar, R. (1978). *A realist theory of science* (2nd ed). Brighton, England: Harvester.

Bhaskar, R. (1989). *Reclaiming reality: A critical introduction to contemporary philosophy*. London: Verso.

Blumer, H. (1956). Sociological analysis and the "variable." *American Sociological Review*, *22*, 683–690.Bohman, J. (1991). *New philosophy of social science*. Cambridge, MA: MIT Press.

Britan, G. (1978). Experimental and contextual models of program evaluation. *Evaluation and Program Planning*, *1*, 229–234.

Bruner, J. (1986). Two modes of thought. In Jerome Bruner, *Actual minds, possible worlds*. Cambridge, MAMA???: Harvard University Press.

Campbell, D. T. (1988). *Methodology and epistemology for social science: Selected papers* (S. Overman, Ed.). Chicago: University of Chicago Press.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge, England: Cambridge University Press.

Cartwright, N. (2000). An empiricist defence of singular causes. In R. Teichmann (Ed.), *Logic, cause, and action: Essays in honour of Elizabeth Anscombe* (pp. 47–58). Cambridge, England: Cambridge University Press.

Cook, T. (2002). Randomized experiments in education: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, *24*(3), 175–199.

Cook, T., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.

Davidson, D. (1967). Causal relations. *Journal of Philosophy*, *64*(21), 691–703.Davidson, D. (1975). Hempel on explaining action. *Erkenntnis*, *10*, 239–253.

Davidson, D. (1980). *Essays on actions and events*. Oxford, England: Oxford University Press.

Davidson, D. (1993). Thinking causes. In J. Heil & A. Mele (Eds.), *Mental causation* (pp. 3–17). Oxford, England: Clarendon Press.

Davidson, D. (1997). Indeterminism and antirealism. In C. B. Kulp (Ed.), *Realism/antirealism and epistemology* (pp. 109–122). Lanham, MD: Rowman and Littlefield.

Dickinson, A., & Shanks, D. (1995). Instrumental action and causal representation. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 5–25). Oxford, England: Clarendon Press.

Ducasse, C. J. (1926). On the nature and the observability of the causal relation. *Journal of Philosophy*, *23*, 57–68.

Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on "scientifically based" education research. *Educational Researcher*, *32*(7), 31–38.

Erickson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher*, *31*(8), 21–24.

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher*, *31*(8), 4–14.

Goldenberg, C. (1992). The limits of expectations: A case for case knowledge of teacher expectancy effects. *American Educational Research Journal*, *29*(3), 517–544.

Gould, S. J. (1989). *Wonderful life: The Burgess Shale and the nature of history*. New York: WW Norton.

Hammersley, M. (1992). Ethnography and realism. In M. Hammersley, *What's wrong with ethnography? Methodological explorations* (pp. 43–56). London: Routledge.

Harre, R. (1975). *Causal powers: A theory of natural necessity*. Oxford, England: Blackwell.

Heil, J., & Mele, A. (Eds.). (1993). *Mental causation*. Oxford, England: Clarendon.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*, 135–175.

House, E. R. (1991). Realism in research. *Educational Researcher*, *20*(6), 2–9, 25.

Howe, K., & Eisenhart, M. (1990). Standards for qualitative (and quantitative) research: A prolegomenon. *Educational Researcher*, *19*(4), 2–9.

Huberman, A. M., & Miles, M. B. (1985). Assessing local causality in qualitative research. In D. N. Berg & K. K. Smith (Eds.), *The self in social inquiry: Researching methods* (pp. 351–381). Thousand Oaks, CA: Sage.

Jankowski, M. S. (1991). *Islands in the street: Gangs and American urban society*. Berkeley: University of California Press.

Johnston, P., & Swift, P. (1994). Effects of randomization on a homeless services initiative: A comment. In K. Conrad (Ed.), *Critically evaluating the role of experiments. New Directions in Program Evaluation, 63*, 67–71.

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.

Kummer, H. (1995). Causal knowledge in animals. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 26–36). Oxford, England: Clarendon Press.

Lather, P. (2004). This IS your father's paradigm: Government intrusion and the case of qualitative research in education. *Qualitative Inquiry 10*(1), 15–34.

MacIntyre, A. (1967). The idea of a social science. *Aristotelian Society Supplement*, *41*, 93–114.
Maxwell, J. (1990). Up from positivism. *Harvard Educational Review*, *60*, 497–501.
Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, *62*, 279–300.
Maxwell, J. (2002). Realism and the role of the researcher in qualitative psychology. In M. Kiegelmann (Ed.), *The role of the researcher in qualitative psychology* (pp. 11–30). Tuebingen, Germany: Verlag Ingeborg Huber.
Maxwell, J. (2004). Re-emergent scientism, postmodernism, and dialogue across differences. *Qualitative Inquiry 10*(1), 35–41.
Maxwell, J. (in press). Using qualitative methods for causal explanation. *Field Methods*.
Maxwell, J., & Loomis, D. (2002). Mixed methods design: An alternative approach. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*, pp. 241–271. Thousand Oaks, CA: Sage.
McGinn, C. (1991). Conceptual causation: Some elementary reflections. *Mind*, *100* (4), 573–586.
Menzel, H. (1978). Meaning: Who needs it? In M. Brenner, P. Marsh, & M. Brenner (Eds.), *The social contexts of method* (pp. 140–171). New York: St. Martin's Press.
Michotte, A. E. (1946). *La perception de la causalite*. Louvain: Editions de l'Institute Superieur de Philosophie. English translation: *The perception of causality*. London: Methuen, 1963.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.
Mohr, L. B. (1982). *Explaining organizational behavior*. San Francisco: Jossey-Bass.
Mohr, L. B. (1996). *The causes of human behavior: Implications for theory and method in the social sciences*. Ann Arbor: University of Michigan Press.
National Research Council (2002). *Scientific research in education*. Washington, DC: National Academy Press.
Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
Premack, D., & Premack, A. J. (1995). Afterword. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 650–654). Oxford, England: Clarendon Press.
Putnam, H. (1990). *Realism with a human face* (J. Conant, Ed.). Cambridge, MA: Harvard University Press.
Putnam, H. (1999). *The threefold cord: Mind, body, and world*. New York: Columbia University Press.
Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: The Free Press.
Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.
Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation* (pp. 3–219). Minneapolis: University of Minnesota Press.
Salmon, W. C. (1998). *Causality and explanation*. New York: Oxford University Press.
Sayer, A. (1992). *Method in social science: A realist approach* (2nd ed.). London: Routledge.
Sayer, A. (2000). *Realism and social science*. London: Sage.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
Shulman, L. S. (1990). *Paradigms and programs*. New York: Macmillan.
Sleeter, C. (2001). Epistemological diversity in research on preservice teacher preparation for historically underserved children. *Review of Research in Education*, *25*, 209–250.
Sperber, D. (1995). Introduction. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. xv–xx). Oxford, England: Clarendon Press.
St. Pierre, E. (2002). "Science" rejects postmodernism. *Educational Researcher*, *31*(8), 25–27.
Strawson, G. (1989). *The secret connexion: Causation, realism, and David Hume*. Oxford, England: Oxford University Press.
Weiss, R. S. (1994). *Learning from strangers: The art and method of qualitative interviewing*. New York: The Free Press.
Yin, R. (1993). *Applications of case study research*. Thousand Oaks, CA: Sage.

AUTHOR

JOSEPH A. MAXWELL is an Associate Professor in the Graduate School of Education, George Mason University, MSN 4B3, 4400 University Drive, Fairfax, VA 22030-4444; jmaxwell@gmu.edu. His research interests include the interrelationships between philosophy, social theory, and methodology; qualitative and mixed-method research design and methods; and the role of diversity in education and other communities.