Abstract

"Qualitative inquiries normally take pains to make clear that they are not generalizable" (Kilbourn, 2006, p. 534). I disagree that this is or should be the norm for qualitative research in education. In this article, I discuss the kinds of generalization that come from qualitative inquiry. I argue that there are several well-established means of in which warranted generalizations can be produced from qualitative research and that the most important of them, theoretical generalization, is of far greater importance to education research and practice than many educational researchers realize.
Introduction

"Qualitative inquiries normally take pains to make clear that they are not generalizable"

(Kilbourn, 2006, p. 534)

Statements like Brent Kilbourn’s that disclaim generalizability for qualitative research are commonplace. Yvonna Lincoln and Egon Guba’s well known 1985 book, *Naturalistic Inquiry*, includes a chapter entitled, “The Only Generalization Is: There is No Generalization.” Harry Wolcott’s in *The Art of Fieldwork* (2005) takes a similar position:

How do you generalize from a qualitative study? [You] might answer candidly and succinctly, “You don’t.” That is a safe and accurate answer. It is the basis on which American anthropology was founded under Franz Boas. With an empiricism directed toward rigorous historical particularism, Boas insisted that no generalizations were warranted from the study of any particular society. (p. 163)

Marilyn Lichtman’s new “user guide” to qualitative research in education (2006) makes a similar claim, stating on p. 7 that qualitative researchers are “not interested in cause and effect or generalizing, but want people to apply [research findings] to their own situations.” And Phil Carspecken in his “theoretical and practical guide” to critical ethnography (1996) writes: “Generalizing across contexts is dangerous” (p. 25).

Statements such as these appear in books and articles intended for novice and expert researchers alike. They come from authors writing about qualitative research in education from the 1980s to the 2000s. They are made by recognized scholars and relative unknowns. Qualitative research, with its reputation for small-scale, researcher-
dependent, and discovery-oriented inquiries, is said to be good for providing detailed
descriptions, identifying relevant factors, and generating plausible hypotheses for more
systematic study. It is not supposed to be good for developing propositions, models or
theories that generalize. Because so many prominent qualitative researchers say so
unequivocally that their work is not generalizable, it is probably safe to say that many
education researchers believe generalizability is irrelevant or unachievable in qualitative
research (see also Schofield, 1990).

But I think they are wrong. Consistent with Kadriye Ercikan’s and Michael
Roth’s argument that “generalization is not a feature of mathematization but a descriptor
for the tendency of inferences to go beyond the context and participants involved in the
research” (2006, p. 22), I make the case that generalizations from qualitative research are
both possible and important.

Types of Generalizations

Probabilistic Generalization

Many people who say that qualitative research is not generalizable seem to define
the concept in probabilistic terms, i.e., as a procedure for making general claims about a
population from a sample, based on statistical probabilities. To generalize in this way,
researchers must provide evidence that the study sample was either randomly or
representatively selected, according to statistical sampling requirements, from the
population to which generalizations will be inferred. If the sampling requirements are
met, then generalizations from the sample to the larger population are said to be
warranted. Robert Yin (2005) calls this approach “statistical generalization.” He writes:
“This method of generalizing is commonly recognized because research investigators
have ready access to quantitative formulas for determining the confidence with which generalizations can be made, depending mostly on the size and internal variation within the [population] and sample” (p. 32). (See other articles in this volume for sophisticated discussions of this kind of generalization.)

Wolcott is one qualitative researcher (of many) who seems to accept this definition. Although he does not explicitly define generalization, he talks about it with reference to averages, frequencies, and distributions (2005, p. 164). In trying to distinguish his own work from the kind that produces generalizations, he aligns himself with the radical particularism of Franz Boas’s style of ethnography, stating that there can be no generalizability from the work of researchers who focus on a site or group chosen by need or convenience, which is often the case with ethnographies.

But despite Wolcott’s strong claim that qualitative research does not produce generalizations, he does not really mean it.1 He subsequently suggests that ethnographic work, even his own, often aspires to a kind of generalization (pp. 163-167). In cases where ethnographers or other qualitative researchers can provide empirical evidence that a particular study site or group is in some sense typical or representative of a larger population, Wolcott suggests that generalizations beyond the particular study may be warranted. For example, if he shows that a Kwakiutl village he studied is similar in important respects to other Kwakiutl villages, or to villages of other groups living along the northwest coast of North America, or even to villages in other parts of the world, then

---

1 There is a difference between saying that one cannot generalize from qualitative studies and saying that one does not intend to generalize from them. Some qualitative researchers claim that they never intend to make generalizations. That is, they eschew theory and comparison and focus solely on the unique and contingent aspects of cases. Schwandt (2001, p. 107) refers to this position as “radical postmodern.” A discussion of this position is beyond the scope of this article.
he believes that findings from his village (despite its being chosen for convenience) can be generalized to the other villages. Borrowing from Margaret Mead, he calls this “approaching generalization” by fitting a particular ethnographic site (however it was selected) into a “larger scheme of things” by which its special characteristics are compared empirically to others. Instead of drawing inferences from a sample to a population based on probabilities, Wolcott draws inferences from a site to others based on observed or reported similarities.

This type of generalization is common in practical reasoning and in education research. Wolcott explains, “I regarded the Kwakiutl village and school of my first fieldwork to be a village and school in certain respects like all other villages and schools, in certain respects like some other villages and their schools, and in certain respects like no other village and its school” (p. 164). To the extent that he can document—usually by reference to previous research--the similarities between his site and others, he makes generalizations from his site to the others. On grounds of similarity, he argues that such generalizations are warranted. Readers may agree or not with his analysis of similarities and thus the appropriateness of his generalizations, but it is disingenuous to suggest that this is not a rational approach to drawing inferences that “go beyond the context and participants involved in the research.” Arguably, this is a more sensible approach to generalization in contexts in which interactions among people, regardless of their overt characteristics as measured in random or representative sampling, can be variable, constructed, and unstable, as is the case in educational contexts.

Other qualitative researchers use more probabilistic ways of approaching generalizability in qualitative research. Some ethnographers and case study researchers,
for example, begin with a sample chosen for its convenience or accessibility. Relying on that sample, they explore initial conditions, identify relevant research questions, narrow the focus, and describe initial results. Then they use surveys to determine whether the initial results hold true for a larger random or representative sample (Holland & Eisenhart, 1990; Schensul, Schensul, & LeCompte, 1999, p. 249). Using this approach, biases inherent in the initial sample but not in larger ones can be identified and eliminated as candidates for generalization.

Consistent with a probabilistic approach to generalization, Joseph Maxwell (2005) makes an important distinction between internal and external generalization in qualitative research. Internal generalization, he argues, is very important, while external generalization often is not. Maxwell writes,

[It] is important to distinguish between what I call “internal” and “external” generalizability. Internal generalizability refers to the generalizability of a conclusion within the setting or group studied, while external generalizability refers to its generalizability beyond that setting or group. Internal generalizability is clearly a key issue for qualitative…studies… If you are studying the patterns of interaction between the teacher and students in a single classroom, your account of that classroom as a whole is seriously jeopardized if you have selectively focused on particular students or kinds of interactions and ignored others.

(Maxwell, p. 115)

Maxwell’s point is well taken. Qualitative researchers in education commonly make internal generalizations of this kind (from a small group of students to the whole class,
from one classroom to a school). When they do, they invoke the requirements of probabilistic generalizability and must add methods such as wider sampling and surveys or evidence of close similarities between sampled and non-sampled groups to justify these internal generalizations.

For Maxwell, external generalizability is another matter:

In contrast, external generalizability is often not a crucial issue for qualitative studies. Indeed, the value of a qualitative study may depend on its lack of external generalizability in the sense of being representative of a larger population…; it may provide an account of a setting or population that is illuminating as an extreme case or “ideal type.” (2005, p. 115)

As noted above, I think statements about qualitative researchers’ disinterest in external generalization are mistaken, but Maxwell’s second point in this excerpt is important, although it is an aside to my main argument. Extreme or ideal cases can reveal what is possible, say when a teacher is particularly successful with normally under-achieving students or when a policy leads to serious unintended consequences in a particular context. Extreme or unusual cases can also illuminate cutting-edge or future-oriented practices, such as teachers who effectively use Web resources in their classrooms. These cases are (presumably) not typical and thus not generalizable. But this does not mean they are bad, useless or weak in rigor. Although not generalizable, special cases are critically important for understanding the variations that develop and the possibilities that exist in educational policy and practice. (See also Schofield, 1990.)
Maxwell’s reservations about external generalizations from qualitative research, like Wolcott’s, stem from the difficulties qualitative researchers can have in meeting the statistical requirements of probabilistic generalization. Maxwell writes,

[A] number of features…lend plausibility to generalizations from case studies or nonrandom samples, including respondents’ own assessments of generalizability, the similarity of dynamics and constraints to other situations, the presumed depth or universality of the phenomenon studied, and corroboration from other studies. All of these characteristics can provide credibility to generalizations from qualitative studies, but none permits the kinds of precise extrapolation of results to defined populations that probability sampling allows. (2005, p. 116)

When generalizations are defined as probabilistic, the people, events, or cases studied must be representative or typical of the larger population for which generalizations are intended or claimed. Qualitative researchers need not strive for this type of generalization in order to produce a worthwhile study, but sometimes they do, and when they do, the standards of probabilistic generalization—that inferences from a sample to a population be justified by statistical or empirical evidence of the sample’s representativeness—apply. Although these standards can be tricky for qualitative researchers to meet, they are not unmanageable. A number of reasonable approaches are common. Qualitative researchers who wish to make generalizations of this type must give careful attention to sampling decisions in the design of their studies. If the intent is to generalize from a qualitative study to ‘what usually happens,’ then the researcher must investigate something that can be documented to be typical or common (Schofield, 1990).
This can be done by selecting a typical site or group in advance, demonstrating the site’s typicality after the fact, or studying multiple sites and identifying commonalities across them.

Thus, probabilistic generalizations—based on statistical probability or contextual similarity, can be and often is produced from qualitative research. And, this is not the only kind of generalization possible from this type of research.

Nomological Generalization

Lincoln and Guba’s (1985) argument that qualitative research does not produce generalizations is based on a nomological definition of the concept. Following Abraham Kaplan, they argue that generalization “must be truly universal, unrestricted as to time and space. It must formulate what is always and everywhere the case, provided only that the appropriate conditions are satisfied” (quoted in Lincoln and Guba, p. 110, from Kaplan, 1964, p. 91). Nomological generalizations are law-like assertions of enduring value that are context-free; any particular phenomenon, sample, data, or example is a special case of such generalizations.

Lincoln and Guba argue, rightly I think, that nomological generalizations are not possible in social science (and perhaps any science), and that qualitative researchers do not accept the reductionist conceit that makes this kind of generalization seem reasonable or helpful. Rejecting nomological generalization and drawing instead on Lee Cronbach’s (1975) notion that the results of social science research are never more than “working hypotheses” for subsequent investigation, Lincoln and Guba propose “transferability,” or the extent of similarity between two contexts, as an alternative to generalization, an alternative that they believe social scientists, including qualitative researchers, can and
should strive for. Transferability of findings or results from one context to another is possible, they argued, if the two contexts are sufficiently similar (p. 124). The extent of similarity suggests the likelihood of transferability. Thus, arguing from different starting points, Lincoln and Guba arrive at a position virtually identical to Wolcott’s and Maxwell’s: Generalizations from qualitative studies based on empirical evidence of similarity across sites and people are both possible and warranted.

Going into more detail than either Wolcott or Maxwell, Lincoln and Guba stress that judgments about transferability depend on detailed knowledge of both sending and receiving contexts. Their perspective is echoed by Frederick Erickson (1986) who writes: “The search is not for abstract universals arrived at by statistical generalizations from a sample to a population, but for concrete universals arrived at by studying a specific case in great detail and then comparing it with other cases studied in great detail” (p. 130). Although a researcher may not be in a position to have such knowledge of both contexts, he or she has the responsibility of suggesting the kinds of contexts to which the results might be transferable and providing sufficient detail about the researched context for a person with intimate knowledge of a second context to judge the likelihood of transferability. Sharan Merriam (1998, p. 211) suggests the term “user generalizability” to refer to another person’s reasons for thinking that a study’s conclusions can be generalized to a context he or she knows well.

Like qualitative researchers who accept the value of probabilistic generalization, those who strive for transferability must think carefully about the selection of a site or sites to study. Sites in which the context can be investigated and described in detail, in which the site can be shown to be typical of other sites, and in which the contexts of the
Generalization from Qualitative Inquiry

sites for generalization can be known are good candidates for qualitative research that is transferable. Methods for coming to know these sites and contexts are many and varied—ethnographies, case studies, histories, surveys, censuses, and previous research all might be useful; coming to know the sites and contexts in detail is what matters in making warranted transfers from one site to another.

Grounded Generalization

Barney Glaser and Anselm Strauss propose another kind of generalization from qualitative research in their book, *The Discovery of Grounded Theory* (1967). Following their constant comparative method, “grounded generalizations” (or theories) are produced as the researcher moves from local situation to local situation, directly following a phenomenon of interest across time and space, investigating everything the phenomena ‘touches,’ describing and interpreting the phenomena in each new situation in terms of the preceding ones, and forming tentative hypotheses that accommodate all previous information and anticipate what new situations will reveal. In this process, the researcher consciously seeks out negative cases, i.e., situations that might force revision or rejection of the emerging hypotheses. The researcher continues this process until new examples of the phenomenon, especially examples that might prove the hypotheses wrong, no longer yield information that is unaccounted for by the hypotheses generated. The final hypotheses become the propositions of a grounded generalization or theory. In this case, the generalization is warranted because it accommodates all the information collected about the phenomenon of interest.

In his early work, Norman Denzin (1978, pp. 191-196) referred to this process of accommodating all the information collected about a phenomenon via the systematic
investigation of negative cases as “analytic induction.” To illustrate the process, Denzin used a study of opiate addiction conducted by Alfred Lindesmith (1947). Lindesmith’s goal was to develop a theory that could explain all cases of opiate addiction. About this approach to generalization, Denzin writes,

This strategy not only forces the careful consideration of all available evidence, both quantitative and qualitative, but makes necessary the intensive analysis of individual cases and the comparisons of certain crucial cases. Thus, Lindesmith did not confine his study only to analysis of individual addicts; he also examined statistical reports on opiate addiction. In addition, he explicitly studied nonaddicts who had regularly received drugs in hospitals in order to isolate the causal conditions present in addiction and absent among nonaddicted hospital patients. (1989, p. 194)

For this type of generalization, a qualitative researcher must give careful attention to the process of the investigation that will lead to a general theory. In this case, selecting a phenomena or site that is typical or one that is accessible enough to expect its context can be well specified is not of primary importance. The researcher intending to develop a grounded generalization can begin his or her work with any instance of the phenomenon but must commit to following it exhaustively through time and space as its connections to other phenomena and in other sites are revealed. This process of generalization is not complete until a theory has been tested against probable negative instances and thus shown to accounts for all known and suspected instances of the phenomenon. I will
return to this type of generalization later in this article, under the heading, Theoretical
Generalization.

Syntheses and Metaanalysis as Generalization

Another approach to generalizing from qualitative research is to develop
techniques for synthesizing the results of qualitative studies about similar topics or
groups. These techniques include case survey method (Yin & Heald, 1975), qualitative
comparative method (Ragin, 1967), and multisite analysis (Miles & Huberman, 1994); all
are strategies for aggregating data across completed qualitative cases. These strategies
consist of steps for locating relevant cases, reviewing the cases for examples of specified
codes or categories, and then identifying patterns in codes or categories that apply (i.e.,
generalize) across the cases. Another strategy, called metaethnography (Noblit & Hare,
1988), involves the identification and cataloguing of the specific concepts, themes, and
metaphors used to report the results of studies of similar topics, e.g., the classroom
implementation of a particular mathematics program. Once the catalogue has been
compiled, the analyst attempts to “translate” concepts, themes and metaphors from one
study into the terms used in another, by asking for example: Are the concepts from the
first study of the mathematics program adequate to handle the concepts of the second and
vice versa? In this manner, more encompassing (generalizing) concepts can often be
identified (see Schofield, 1990, pp. 222-226, for summaries of these approaches to
qualitative synthesis).

Without belaboring the point, I want to note here that I have just discussed six
kinds of generalization that are relevant to and can be obtained from qualitative research:
probabilistic generalization, transferability, user generalization, grounded generalization,
synthetic generalization, and meta-analytic generalization. Perhaps there are others. Let it not be said any longer that qualitative researchers don’t “do” generalization. If generalization is a concern for qualitative researchers (and often it is) or for those who evaluate qualitative research, there are a number of established ways it has been and can be addressed. In the next section, I want to elaborate on one other type, theoretical or analytical generalization, a form that is arguably more important to qualitative researchers and to education research than any discussed so far.

Theoretical Generalization

Although Maxwell is concerned about probabilistic issues, he is careful to point out that generalizability from qualitative studies is more often based on the development of a theory that can be extended to other cases or refined in light of them (2005, p. 116). Robert Yin calls this approach “analytic generalization” (2005, pp. 10, 37-38). Charlotte Davies refers to it as “theoretical inference”—where “the conclusions of [a qualitative study] are seen to be generalizable in the context of a particular theoretical debate rather than being primarily concerned to extend them to a larger collectivity” (1999, p. 91). Davies uses the example of Cynthia Cockburn’s, *In the Ways of Women: Men’s Resistance to Sex Equality in Organizations* (1991), an ethnographic study of the introduction of equal opportunity policies in four British organizations, to illustrate the difference between what she calls “empirical generalizations” (similar to Wolcott’s “approaching generalizations”) and theoretical inference. Davies writes:

[Cockburn] offers some empirical generalization in that [her conclusion is] not restricted to the specific four organizations she studied but is meant to be applicable to other similar organizations in British
society, and, perhaps with some modifications, to other Western industrial societies. On the other hand, her more significant generalizations [i.e., her theoretical inferences] have to do with the forms of resistance both formal and informal that characterize the introduction of [equal opportunity] policies. Such generalizations are likely to be of much greater explanatory value in quite disparate situations… This sort of generalization relies upon a case-study method in a very different way than as a representative of a class of cases. (pp. 91-92)

In Cockburn’s work, cases (ethnographies, case studies, etc.) are used to develop and then refine emerging theoretical inferences about the nature and process of resistance to equal opportunity policies. Davies continues:

[This sort of generalization] proceeds by a gradual accumulation and ‘constant comparison’ (Glaser and Strauss, 1967) of cases in which, rather than seeking to show repeated instances of particular conjunctures of occurrences leading to a predictive causal statement, the ethnographer actively seeks the differences and variations whose explanation will refine, strengthen and make more profound the developing explanations that constitute valid generalization in ethnographic research. (p. 91)

In striving for theoretical generalization, the selection of a group or site to study is made based on the likelihood that the case will reveal something new and different, and that once this new phenomenon is theorized, additional cases will expose differences or variations that test its generalizability. The criterion for selecting cases from which one
will generalize is not random or representative sampling but the extent to which the cases selected are likely to establish, refine or refute a theory (see also Schwandt, 2001, p. 106).

As Davies indicates, the process of theoretical generalization proceeds in a manner similar to the process of grounded generalization outlined by Glaser and Strauss. The difference is that the goal of grounded generalization is to produce new theories or explanations whereas the goal of theoretical generalization is to make existing theories more refined and incisive. It seems to me that particularly strong qualitative research programs would aspire to both generalizing goals—first to develop a grounded theory and then to refine it by extension to other cases.²

Howard Becker (1990, p. 238) describes theoretical generalization as the attempt to develop a refined understanding of a generic process, such as the functioning of Erving Goffman’s “total institutions,” that have wide applicability in social life. Becker explains that the point of theoretical generalization is not to show that every site with the characteristics of a total institution produces the same results, but rather to show how each new site potentially represents different values of a generic process. In other words, a theoretical generalization can be true although the results in specific cases are different. Becker provides an example from studies of men’s prisons (in the 1960s and 70s) and the attempt to generalize their results to a prison for women:

Students of prisons…had demonstrated that, in the men’s prisons they studied, inmates developed an elaborate culture. They created a convict government that took over many of the functions of keeping order in the joint; they developed quasi-markets in cigarettes, drugs, tailor-made...

---

² Many education researchers doing qualitative work claim that they are doing “grounded theory research.” However, very few follow the methodological steps outlined by Glaser and Strauss and even fewer ever produce a grounded theory.
clothing, and a variety of personal services; they organized sexual activity; they enforced a strict code of convict behavior emphasizing the necessity of never giving information about other prisoners to prison guards and officials.

Analysts of prison culture attributed these inventions to the deprivations of prison life... The generalization was, prisoners develop a culture that solves the problems created by the deprivations of prison life.

... [Other researchers], with this theory in mind, studied a women's prison. They didn’t find any of that. Quite the opposite. Even the officials of the prison complained about the lack of a convict code: The women were forever snitching on one another in a way that made a lot of trouble. There was no underground market in much of anything. Sex life was not organized in the predatory style of the men’s prison; instead, the women developed pseudo-families, with butches acting as the husbands and fathers of a collection of wives and daughters.

Do these differences invalidate the generalization that the deprivations of prison life lead to the creation of a prison culture?... Not at all... [T]he theory wasn’t wrong, but you had to put in the right values of the variables to see how it was right. You could still say that the deprivations of prison life led to the creation of prison culture, but that this was true only if you understood that prison deprived women of different things than men. Women were not deprived of autonomy because, on their own testimony, they had never had it; they had always lived under
the protection of a man—a father, husband, or lover. They were, however, deprived of exactly that kind of protection. So they didn’t develop a convict government, but they did develop a system of homosexual relationships in which one woman stood in as the masculine protector…

In short, women are deprived of different things, both because their lives on the outside and, therefore, their needs on the inside differ, and because the prison is run differently for them. Their culture responds to that difference. The generalization is still true, even though the results are quite different. (pp. 240-241)

Note that Becker is not suggesting that there are no better generalizations, or even that this generalization will endure over time. He is suggesting, however, that until a better generalization is proposed and shown to be warranted, this one accounts for multiple, even disparate cases.

Becker’s view of theoretical generalization from qualitative research is similar to Clifford Geertz’s. In discussing cultural theory, Geertz writes:

What generality it contrives to achieve grows out of the delicacy of its distinctions, not the sweep of its abstractions. ...Studies do build on other studies, not in the sense that they take up where the others leave off, but in the sense that, better informed and better conceptualized, they plunge more deeply into the same things… One can, and this in fact is how the field progresses conceptually, take a line of theoretical attack developed in connection with one exercise in ethnographic interpretation and employ it
in another, pushing it forward to greater precision and broader relevance.

(1973, pp. 25-26)

Geertz continues,

…the theoretical framework in terms of which such an interpretation is made must be capable of continuing to yield defensible interpretations as new phenomena swim into view…If they cease being useful with respect to such [phenomena], they tend to stop being used and are more or less abandoned. If they continue being useful, throwing up new understanding, they are further elaborated and go on being used. (1973, pp. 26-27)

There are many examples of emergent theoretical generalizations in existing qualitative studies in education research, and some of them appear to be strikingly enduring. Janet Schofield’s study of a desegregating school (1989) in the U.S. provides a case in point. About this study, Schofield writes:

After I observed extensively in varied areas of the school and interviewed a large number of students, it became apparent that the white children perceived blacks as something of a threat to their physical selves. Specifically, they complained about what they perceived as black roughness or aggressiveness… In contrast, the black students perceived whites as a threat to their social selves. They complained about being ignored, avoided, and being treated as inferior by whites, whom they perceived to be stuck-up and prejudiced… Such findings appear to me to be linked to the black and white students’ situation in the larger society
and to powerful historical and economic forces, not to special aspects of
the school [she studied]. The consequences of these rather asymmetrical
concerns may well play themselves out differently in different kinds of
schools, but the existence of these rather different but deeply held
concerns may well be widespread. (p. 221)

Although Schofield’s school ethnography was conducted in the 1970s, her conclusion
that asymmetrical concerns differentiate the reactions of students from different racial
groups in school and thereby maintain social distance between them is almost certainly
ture in U.S. schools today. Although particular concerns will vary from group to group,
school to school and over time, the idea of asymmetry in racial group members’ concerns
about other groups is likely a pervasive pattern (a generic process?) in contemporary U.S.
schooling. Had this idea been systematically pursued in subsequent studies of schools, it
might have been refined and developed into a theory of inter-group relations in school
with broad generalizability to the U.S. context.

Signithia Fordham’s book, *Blacked Out* (1996), provides another example. She
describes in detail how and why black students in a Washington, DC, high school
struggle with the meaning of school work and achievement. Again, the particular actions
and beliefs of the students in her study may not appear elsewhere, but the forms of
resistance to schooling that she identified are likely to be widely applicable to minority
describes how Mexican-American students in Tucson make sense of using both English
and Spanish in their lives. Here again, the particular characteristics of the students
probably do not generalize, but the emotional loading of language use that she identified
is likely relevant to bilingual education wherever it takes place. Mary Metz (1990) describes an ideology of schooling that she discovered in a study of several Midwestern high schools in the 1980s. At that time she found a common script for a legitimate U.S. school (which she referred to as “The Real School”); today, thirty years later, that same script is evident in NCLB and other current policy initiatives (Metz, 2005). These patterns—of resistance, language loading, and taken-for-granted understandings of what makes a “real school”—are the kinds of (potentially) powerful theoretical generalizations that can result from qualitative research in education. They represent in some detail descriptions that capture and recurring processes that explain (in part) major educational problems, including low achievement, weak second language skills, and sometimes excessive discipline, in the U.S. education context.

A discouraging fact about education research in the U.S. is that few qualitative researchers pursue investigations designed to test or extend the theoretical generalizability of other qualitative researchers’ conclusions. Few even attempt this with their own work. Emergent mid-range theoretical generalizations such as those exemplified in the research just discussed would seem to be prime topics for qualitative education researchers who are interested in generalizability to pursue: Each generalization is based on extended and in-depth qualitative research in a particular local context, yet the conclusions from each constitute a plausible explanation for a pervasive phenomenon of U.S. schools and a point of departure for additional case studies and theoretical development. It is surprising (to me) that studies to extend the theoretical generalizability of qualitative research are so rare, especially when the contribution of such work to the field of education seems so great. Perhaps this lapse is in part a
consequence of misguided discourse suggesting that one cannot generalize from qualitative research.

Theoretical generalization from qualitative studies also has been discussed by researchers with critical and postmodern perspectives. Although profoundly skeptical of the way in which generalizations can be used to elide differences and promote essentializing stereotypes, these researchers nonetheless suggest that reconstructive analyses of multi-voiced texts can illuminate theories at work in the social and communicative connections that constitute everyday life (Carspecken, 1996; Denzin, 1989; Lather, 2003). Similar to mid-range theories of generic educational phenomena, mid-range theories of constitutive relations in narrative, social interaction, and communication are areas of tremendous potential for theoretical generalization from qualitative research.

Conclusion

In 1990, Janet Schofield argued that there was a consensus emerging among qualitative researchers about generalizability. She wrote:

First of all, there is broad agreement that generalizability in the sense of producing laws that apply universally is not a useful standard or goal for qualitative research. In fact, most qualitative researchers would join Cronbach (1982) in arguing that this is not a useful or obtainable goal for any kind of research in the social sciences. Second, most researchers writing on generalizability in the qualitative tradition agree that their rejection of generalizability as a search for broadly applicable laws is not a
rejection of the idea that studies in one situation can be used to speak to or
to help form a judgment about other situations. (p. 208)

If this consensus was emerging in 1990, it seems to have dissipated at least
rhetorically. The conventional “wisdom” about qualitative research now seems to
be that it cannot be generalized. In this article, I have argued that this view is
wrong and misleading. Even those people whose writings have been used to
establish this view do not necessarily believe or practice it.

For educational researchers, most of whom want to improve education or
understandings of education, it hardly seems likely that they would willingly
devote their professional time and energy to studies that do not generalize in some
way. Who really does not want to apply or infer from what has been learned in a
study beyond the particular case? Fortunately, there are numerous, well-
established ways of approaching generalization from qualitative research. In my
view, theoretical generalization is an especially promising avenue for future
qualitative research.
References


