A pluralist view of generalization in qualitative research

Staffan Larsson

[a] Department of Behavioural Sciences and Learning, Linköping
University, Linköping, Sweden
Version of record first published: 18 Mar 2009.

PLEASE SCROLL DOWN FOR ARTICLE

A pluralist view of generalization in qualitative research

Staffan Larsson*

Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

A common way of discussing generalisation is to search for one conception – a monist view. Another approach is to create a dichotomy between quantitative and qualitative research, each having a single definition – a dualist perspective. A pluralist view is argued for here, i.e. the existence of several lines of reasoning, each of which can be evaluated in specific cases in terms of usefulness, strengths and weaknesses. Five different lines of reasoning about generalization, which could possibly be useful in qualitative research, are presented. They are: 'studies that undermine established universal "truths"', 'the ideographic study', 'enhancing the generalisation potential by maximizing variation', 'generalisation through context similarity' and 'generalisation through recognition of patterns.' Each is critically commented on.

Keywords: generalization; ethnography; qualitative methods; human science; methodology

Generalization as an issue

Texts on qualitative methods often demonstrate a lack of enthusiasm for the problem of generalizations. The limited number of elaborated texts on the topic is one sign. Sometimes the word generalization is used in paradigmatic struggles – as friend or foe. The concept can be dismissed in social science by some researchers, e.g. Usher (1996, 14). Scheurich declares that his postmodernist orientation explicitly links the word to a modernist realism: 'While these generalisations are said to represent reality, in my mind they mostly represent the mindset of the researcher' (1997, 64). This kind of dismissal operates on the assumption that the word generalization has a clear, singular meaning. Scheurich is in fact making a generalization, while at the same time rejecting the use of it. It illustrates the difficulty of escaping the 'mindset', which is signified by the word. And his generalization on the epistemological position of researchers who find the word useful is not convincing. Greenwood and Levin are, on the other hand, critical of the universal critique of generalization, being 'without any sense of social or moral responsibility' (2005, 55). Social responsibility might be an argument, but at the heart of the matter are adequate conceptualizations of generalization as a phenomenon.

Instead of marginalizing the issue, I will argue for the need of an elaborated discourse on generalization. Furthermore, a case is made for the necessity of thinking in terms of several lines of reasoning on generalization. A third thesis is that different

---

*Email: staffan.larsson@liu.se

kinds of research are faced with particular problems of generalization and have to use appropriate lines of reasoning on the issue of generalization.

This article discerns a number of such qualitatively different lines of reasoning beyond a crude commonsense singular meaning to signify the phenomenon. Awareness of a repertoire of possible lines of reasoning will hopefully make arguments about generalization more precise and rational: you need a number of tools, in order to be able to choose what is appropriate for various tasks. The need for an elaborated discussion is especially strong in educational research, since it covers such a wide variety of research traditions and methodologies.

Words used and what they signify

However, the issue is complicated: the word generalization 'competes' with other words, which seem to have at least a family resemblance, e.g. transferability (Lincoln and Guba 1999) or external validity (Campbell and Stanley 1963). How to delimit the phenomenon is not obvious, since the words used also indicate slight differences in what phenomenon we are talking about. Our trouble with exact meaning might be allayed by Wittgenstein's remark on the lack of exact concepts, when he points out the usefulness of a rough indication of something: 'If I tell someone "Stand roughly here" – may not this explanation work perfectly?' (1958, 41). I suggest that we accept a rough, but useful, point of departure in order to arrive later at a more elaborated, still not exact, but more useful understanding.

What is indicated by words like 'transferability' or 'generalization' is fundamentally part of everyday life. Our language is permeated by taken-for-granted generalizations. People generalize, researchers generalize and even the postmodernists quoted above generalize. Readers of any kind of study often ask: Where and when are these interpretations useful? Questions about generalizability seem to remain topical not only for researchers, but often even more so for the interested general public. Educational researchers often express hopes of contributing to public debates. If they content themselves with a denial of the need to answer questions about the potentials and limits of the claims, they might not be taken seriously. To deliberate on this phenomenon instead of dismissing it can therefore be a good idea – even if concepts are only 'usable'.

Monist, dualist or pluralist view

In order to elaborate on the issue we need a rough starting point. By means of a Latin dictionary we can catch a glimpse of an old root. The Latin word 'gener' is given the meaning 'relative through marriage' and 'generalis' refers to 'that which is of a certain family or kind' alternatively 'that which includes everyone in a specific family, common' (Ahlberg, Lundqvist, and Sörbom 1964, 375, author's translation from Swedish). The meaning in our contemporary discourse is different, but we can see a family resemblance to these ancient uses of the word, which gave birth to the contemporary expression 'generalization'. However, the first challenge is the tendency to look for a single meaning.

It is not only in Usher's and Scheurich's dismissal that we can observe the meaning of generalization as being understood as uniform – one will find the same in a dictionary of statistics and methodology defining generalization (Vogt 1999, 121). Being on opposite ends of possible positions on many aspects of methodology, they share a monist view on generalization. As I indicated before – there is something sound in looking for

such a common ground. However, it should be considered as, at best, preliminary since the monist understanding of generalization will create difficulties. The most funda-mental flaw is when the concept gives a particular aspect of generalization a universal position. A view based on representation through strict sampling from a defined population is limited to a certain kind of research and is, for instance, not very useful in qualitative research. Nor is it useful in classic experiments – laboratory experiments will routinely be met with some scepticism, until they are replicated in other labora-tories. Even very ambitious large-scale experiments or quasi-experiments will face the same question from sceptics concerning other populations from which the sample was drawn (Shulman 1997, 14). Qualitative researchers sometimes borrow from the vocab-ulary of statistics. In their classic qualitative study of one life history, Thomas and Znaniecki use the metaphor of representativeness, claiming this life history to be repre-sentative of the mass (Chase 2005, 667). Glaser and Strauss (1967) use the notion 'theo-retical sampling' and thus also borrow from descriptive statistics. However, the common denominator is very superficial here. Consequently, even before discussing qualitative research more in depth, we can sense the difficulties a monist position faces when we want to move beyond our rough point of departure.

However, most of the discussion on the issue among qualitative researchers oper-ates with various dualistic divisions. In the 1994 edition of *Handbook of Qualitative Research*, Denzin and Lincoln propose that researchers adhering to constructivist or postmodern convictions use the word 'transferability'. They seem to suggest a dichot-omization of the issue, each with its own line of reasoning according to the epistemol-ogy chosen. Here, we can discern one kind of dualistic answer. On a certain level, it makes sense – we need more distinctions, not fewer. However, the success of this neologism seems to be limited. The term 'transferability' is not used by Guba and Lincoln in a chapter about existing paradigms in the 2005 edition of the same hand-book, while 'generalization' is used in relation to critical theory (Guba and Lincoln 2005, 194, 196). However, the need for various distinctions seems to have survived.

Hammersley has argued for a qualitative difference between empirical and theo-retical generalizations and that ethnographic work can use both forms. The empirical kind is about claims that 'the particular setting investigated is typical of some larger whole or aggregate' (1992, 86). Theoretical generalization refers to 'drawing conclu-sions about one or more social scientific theories from the features of the local events they observe and describe' (91). He defines theoretical as 'statements about necessary relationships among categories of phenomena'. Hammersley thus provides a different dualism, dividing ethnographic and social science into two parts in this respect. He does not underscore a qualitative/quantitative dividing line but, rather, emphasizes the support of survey research in making empirical generalization claims for ethnographic research.

Schofield (1993) also creates a dualism in her discussion on generalization. Here, it is a dualism between qualitative and quantitative research. She argues that general-ization from a case can be made by considering the similarities of the context between a researched case and the cases that can be generalized to.

As we have seen, there are a number of ways to discern dualistic divisions concerning the meaning of generalization. Both monist and dualist answers underesti-mate the complexity of the problem, hopefully which this text will show. Rather, there are good reasons for operating with a plurality, discerning not only arguments for different ways of doing generalization, but also for cases when generalization with good reason is not an important issue. A plurality of meanings is needed in order to

address a variation of *generalization problems*. Here, five qualitatively different lines of reasoning in qualitative research are suggested. I do not claim these five to be the only possible, but they should at least illustrate my arguments for a plurality of lines of reasoning. Hopefully, they will also point to the dialectic between the specific research design and what will become a warranted reasoning around generalization – different research designs having different kinds of generalization problems.

Five lines of reasoning on generalization

Five possible lines of reasoning are sketched. The first two argue that there is no need for generalization, i.e. certain kinds of empirical research are meaningful without any claims of generalization. The other three argue in different ways in favour of possible ways of generalizing. My point of departure is very pragmatic insofar as I try to think of the practice of qualitative research with all its diversity and pose the question: What are the possible answers to a question about generalization in different cases? I am looking for the fundamental logic in ways of answering such a question in the social practice that we call qualitative research. I am not looking for philosophical justifications but, rather, for how people use interpretations that emanate from research. The first two lines of reasoning concern cases when generalization claims are redundant or not appropriate:

(1) The ideographic study;
(2) Studies that undermine established universal 'truths'.

The next three can be useful when generalization is called for:

(3) Enhancing generalization potential by maximizing variation;
(4) Generalization through context similarity; and
(5) Generalization through recognition of patterns.

Two kinds of research that make generalization claims redundant

Here, I sketch two different lines of reasoning, each completely different, but which still end up in disqualifying the necessity of making generalizations. However, even if generalization is redundant, the lines of reasoning are not redundant as they are important in order to warrant the lack of generalization claims in specific pieces of research.

The ideographic study

In this case, generalization is not meaningful because these studies belong to a particular kind of study where the underpinning logic is one of the unique pieces put together in patterns. One can use the metaphor of a jigsaw puzzle. The role of a specific study is not to say something about other contexts, but to contribute to the broader picture by filling a 'hole' in the whole. Let me exemplify. An investigation of the conceptual inspiration of Paulo Freire's thinking would be justified even if you could not generalize. The study will be a contribution to the understanding of the tapestry of educational philosophy, e.g. how Freire's views connect to various

thoughts in philosophy generally. It is easy to find examples also from research on the history of educational reforms as unique historical situations. The term ideographic has been coined for this kind of research by Windelband – a German anti-positivist philosopher at the end of the nineteenth century (von Wright 1971, 5). The opposite was nomothetic, where research aims at discovering universal laws that must be generally applicable. Generalization is in principle redundant in the case of ideographic studies, since such studies focus on: 'the descriptive study of individuality' (von Wright 1971, 5). This late nineteenth century anti-positivism rejected methodological monism, which was argued for by advocates of the positivistic philosophy of science. The celebration of the ideographic was part of a revival of hermeneutics as a method for human science. In hermeneutics, the focus is on understanding the specific case as a part of a larger pattern, i.e. each case has a specific role in the larger pattern, like pieces in a jigsaw puzzle. In history, it is easy to see how investigations into certain events are motivated, not because they represent similar events, but because each event is linked with other but different events. Talk about generalization in this context is thus a mistake, since it would be a lack of understanding of the fundamental logic of such research. However, I do not think that all kinds of qualitative research have this logic. In spite of this disagreement, these late nineteenth century philosophers' way of portraying the practice of research is useful in relation to a considerable number of cases. It is therefore perfectly legitimate to investigate the emergence of educational systems in a specific country, without any aim of generalizing the interpretation, like Durkheim's (1985) work about the development of secondary education in France. Even though historical studies or biographies could be seen as prototypical of the first line of reasoning, there are certainly historical investigations that are based on a logic where generalization problems are at the forefront and therefore follow a different kind of logic. Social history, when the purpose is to describe general patterns among the population based on limited information about a number of persons, is one example. We can also think about biographies, which trigger questions about generalization, e.g. a schoolteacher representing others of his kind.

Sometimes there are studies, where there are parts, where generalization is not called for and other parts, where questions of generalization arise. Policy ethnographies can be an example. Generalizing from the production of a specific political agenda on education in the EU administration might not be necessary, but when the consequences on the local level are studied, the issue of generalization will be brought up.

Critical comments

One difficulty with this line of reasoning is that the significance of the case becomes crucial. The philosopher or the reform must be important in some way for the broader context. The significance has to be convincingly argued for. The focus on the significance of the single case places certain phenomena in the foreground and others in the background. It is easy to argue that decision-making among elites in central positions are important single cases, but it is difficult when the focus is on the local level. Another, similar, effect is that it is easy to argue that dramatic events are significant cases, while changes in everyday routines are less easy to use as important single cases, where generalization is not needed. So this line of reasoning is limited in use and can only be applied where appropriate, i.e. the German philosophers' claims were a contribution, but when they argued that there was only one alternative to the logic of natural science, they overrated the ideographic notion.

Studies that undermine established universal 'truths'

In this case, the situation is quite different from the first line of reasoning. The focus here is on 'negative' cases, i.e. where the research creates doubt about something, which is generally thought as true. We can call these latter convictions universal claims, since they take the position of being taken for granted or are actually the established view within a discipline. Each single case that deviates from the established truth will be of vital importance, since it falsifies one essential aspect – the universalistic claim. After such a study has created doubt, the reach of a claim is reduced. This way of arguing can be found in post-structural approaches. Söndergaard describes the ambition to 'contradict the obvious, to think against the stream of what is taken for granted' (2002, 1, 91). Cases that break the rule, which are not in accordance with the available discourses, will do the job of troubling or destabilizing the taken-for-granted. What seemed to be essential is contingent, i.e. true in some cases, but not always. The undermining effect is enough to make the research meaningful. The line of reasoning here goes like this: all cases must be included if something is universal, i.e. any deviant case undermines the universalism. This does not mean that the numbers *per se* are not interesting and meaningful but it is a completely different question, irrelevant in relation to the claim of universalism. An example: in a study by Brown, Cervero, and Johnson-Bailey (2000), we can discern an undermining effect on the idea that it is progressive to have a pedagogy, which gives students a strong position in the classroom. They describe how black, female math teachers' authority is challenged by students. Students' image of a 'genuine' math teacher is a white male, i.e. their challenge is based on race and gender constructions and the conclusion will undermine a dominating idea about the political implications of a pedagogy where students are influential. This case shows that such a pedagogy could support racism and sexism. Phoenix (2004) presents findings, undermining critical pedagogy's claim to liberate from oppression, showing how such pedagogy operates in relation to black students' views of masculinity. That universal claims about human nature have been undermined by case studies is nothing new. Mead's (1963) work on cultural variation of gender constructions is a classic example from social anthropology, producing doubt about that something is 'natural'. Undermining universal claims when it is possible is extremely important since false universal claims are used to delimit human understanding of the possible range of being – i.e. they suppress possibilities and place false limits on freedom. History is full of examples where science has presented arguments about human nature that have been used to legitimate racism, sexism, etc.

It is easy to see it as parallel with Popper's focus on falsifying as the central attitude in research. However, the underlying assumption is, in his case, that science is about universals, however preliminary. Social or human science has such a striving towards universals only to a limited degree. We probably meet more universalisms in the public debate, where citizens' opinions are influenced.

Critical comments

This line of reasoning is limited insofar as it presupposes that there are universal claims to investigate. In contemporary human and social science, such grand ambitions are often not put forward. To some extent, this is based on the fact that these parts of academia view human action as socially and culturally constructed and thus do not follow universal laws. However, humans are also the legitimate objects of

natural science, as well as social or human science, which create a fruitful area of debate on questions about human nature.

When generalization claims are called for: three lines of reasoning

I would argue that qualitative studies, with the above exceptions in mind, on the whole have difficulties in avoiding making claims about generalization. The phrase 'one cannot draw any conclusions about any other situation than the investigated cases' is sometimes used in defence of qualitative studies. Taken seriously, it will reduce the interest in many qualitative studies to practically nothing. If someone has made a study of a classroom in the spring of 2005, it is difficult to take seriously if there are no ambitions to say something that can be of use outside this situation in time and space and the persons involved: '… there must be a *capacity* for generalization; otherwise there would be no point to giving such careful attention to the single case' (Wolcott 1994, 113). I want to show possible solutions, which can be used, when appropriate, adding three qualitatively different lines of reasoning to the two already presented.

Enhancing the generalization potential by maximizing variation

This line of reasoning is useful in studies where a number of cases constitute the empirical basis. Qualitative interview studies are examples. This logic is based on the fact that a sample is drawn, but the logic of sampling is opposite to the standard of statistical sampling. Instead of relying on random chance in order to calculate the representativeness, one wants to cover a variation of qualitatively different cases of a phenomenon. To understand how pupils experience the teaching about religions, one might get a fuller understanding by choosing pupils with as mixed religious background as possible instead of a representative sample. Covering more of the variation in qualitative different views will enhance the generalizability of the study. In an interview study based on random sampling, the most common answer would be really well represented and unusual answers would be less or not at all represented. In order to maximize the differences, a sample should be based on qualified guesses about how to achieve this broad variation. This thought has a certain similarity with 'theoretical sampling' (Glaser and Strauss 1967): Sampling should be based on what was already known and what was needed next in order to push the understanding of the researched phenomenon further. The variation of a phenomenon should be investigated. The idea of maximizing variation has been most explicitly expressed within the tradition called phenomenography (Marton 1994), where the focus is on describing variation in ways of seeing a phenomenon. It is often used in studies of learning, e.g. in science education. In order to generalize from a certain study, one needs to optimize the probability that as many qualitatively different cases or categories as possible will be possible to describe. This means that the uncommon case is as important as the most common kind of case. In this kind of reasoning, one cannot generalize from one specific category or case but only from the whole set – the variation in the study should be expected to exist also in relevant situations that one wants to generalize to. If a wise selection of persons or cases, which could be expected to be diverse, has been made, one could expect to have covered the variation relatively well. Kennedy summarizes this line of reasoning: 'The range of characteristics included in a sample increases the range of population characteristics to which generalization is possible' (1979, 665).

Critical comments

The usefulness of this line of reasoning is limited to such studies where one operates with not too few cases, for instance, qualitative interviews. It is really not an option for traditional case studies, where concentration on one or two cases is often recommended in order to conduct an in-depth study. Another problem is the choice of cases or persons: it is not easy to predict the real difference on the basis of surface impression or formal characteristics. This presupposes a deterministic logic, which is often not realistic – cases often turn out not to be what they looked like or persons with certain social characteristics do not follow prejudices about how such persons should think. A third problem is the lack of knowledge about the real breadth of the variation: in terms of generalization, it is impossible to know how many undetected variants there are in real life.

Generalization through context similarity

The similarity between a researched context and other contexts are focused on in some texts on generalization. Thorne's (1993) study of how gender identities are shaped by everyday interaction in schoolyards should, according to this logic, be possible to generalize to similar schoolyards. Schofield (1993) argues that this conception should be the foundation of a view of generalization that is adapted to the essence of qualitative research. She refers to a number of authors who have developed such arguments about context similarity. The concept of 'transferability' has also been defined as similarity between contexts – Lincoln and Guba define the constructivist or postmodern version as something that is related to abstracted results: 'the transferability of which is an empirical matter, depending on the degree of similarity between sending and receiving contexts' (1999, 404). The centre of attention here is obviously on context and on the similarity between contexts. Strauss and Corbin suggest that 'explanatory power' should replace generalizability. In this case, prediction is a key factor: 'Therefore, in writing the theoretical formulations that evolved from our study, we specify the conditions that give rise to certain phenomena' (1990, 267). One is given the impression that they operate with causality as a presupposition, i.e. if conditions are known, the consequences should be possible to predict. I judge this to be a variant of context similarity, even if they do not express it in such a way. Hammersley (1992), whose line of reasoning about empirical generalizations is about similarity of settings, views similarity as an open question that has to be answered by means of empirical support, e.g. survey data. Lincoln and Guba point out that judgements about generalizability presuppose that contexts are known, which often means that those who want to use the research are better able to judge than the researchers. Consequently, they draw the conclusion that 'the responsibility of the original investigator ends in providing sufficient descriptive data to make such similarity judgements possible' (1999, 404). Here, I want to stress that something quite different, which has important consequences, is introduced. When they argue that the researcher is not necessarily the person with the obligation to judge the generalizability, they change the power relation between the researcher and his audience. It is the audience that is often in the best position to judge the similarity of a context with the one portrayed in the research work. The role of the researcher then changes into one where the description of the context of the interpretations is given this new function: to communicate a context to an audience, which has the role of judging whether some context they know about is similar

to the researched context. They refer to Geertz's (1999) reasoning about 'thick description' as a foundation for drawing conclusions about context similarity. Geertz's text discusses the validity of descriptions in terms of thin and thick: thin is a description that only describes behaviour while a thick description also includes actors' interpretation of the world. A wink of the eye could be seen as something that has a meaning (thick) or as a reflex without any message (thin). If I interpret Lincoln and Guba with the help of Geertz's text, it could be concluded that the concept 'context' should not only refer to descriptions of material circumstances and actors' behaviour but also to the interpretational world, e.g. a similar culture.

Critical comments

Descriptive statistics' way of dealing with generalization is based on probability theory. This is not an option in qualitative research, which is normally based on one case or a limited number of interviews. Instead, the focus must be on what is empirically known rather than theoretical assumptions. The focus on similarity between the research context and other similar contexts becomes a kind of parallel to the relation between sample and population. Instead of operating with a similarity that is an *a priori* assumption, as in traditional sampling, one must judge the similarity empirically and *a posteriori*. One problem here is the difficulties in judging when a similarity is present. It is obvious that the varying exactness in describing a case points to varying numbers of similar cases. The exactness of similarity searched for in order to draw conclusions about unresearched cases often seems to be unknown to both the researcher and the reader of a study. Another problem concerns a more theoretical aspect. Generalization via context similarity presupposes that the context determines the phenomenon or pattern. I.e. an idea that a specific context will always hold the same qualities in a phenomenon or pattern. It is not difficult to imagine cases where persons act differently in the same context or even that the same person acts differently in the same context on different occasions. The assumption that qualities are determined is, to say the least, an underestimation of the complexities of human action. One is reminded of the fact that the same person is able to operate with different interpretations of the same phenomenon. Different persons act differently in the same context, e.g., because of different personal histories. Normally, it is not practical to check such things in order to judge the possibility of drawing conclusions about generalization. One limitation as regards this conception is that it is suited to case studies where there is an abundance of context data. It is less suited to, for instance, qualitative interviews where context data are less prominent.

Generalization through recognition of patterns

Research texts can communicate ways of seeing something, often with the ambition to transcend old or taken-for-granted ways of understanding the studied phenomena – this is the 'heuristic validity' of interpretational research (Larsson 2005). Qualitative research often produces such interpretations – theoretical constructions, concepts or descriptions, i.e. patterns or configurations, which can be recognized in the empirical world. The reader is invited to notice something they did not see before. We can view this as a variant of generalization, the communicated pattern is recognized in new cases. An example, Gamble (2001) describes the 'pedagogy' of a South African furniture carpenter introducing a novice into the trade in wordless communication. Her

description cannot be generalized in a simple way to all other carpenters introducing novices, but knowing Gamble's description and interpretation of the case, one may recognize similar wordless pedagogy in teaching, which can be similar, but also different in varying degrees from the original case. Since in this case it is a process, not a person or a context, that is in focus of her research, it is very difficult to predict when or where something similar will happen. On the other hand, we can be alert to the potential use of Gamble's study when we are observing situations and persons, which is roughly similar. But such context similarity is only a potential for recognizing the process. There is a loose relation between process and context. Another familiar example, in a study of a North American university, a 'hidden curriculum' is described (Snyder 1971). With the help of the analysis of the hidden curriculum and the description of it, it becomes possible to recognize similar patterns in other educational institutions. This pattern has reached a huge audience at least indirectly, i.e. many can recognize the pattern of a hidden curriculum in various contexts. However, it would be too much to expect such a description or interpretation to be appropriate in all other similar contexts – it is only potentially useful in other cases. The line of reasoning here is that generalization is about the potential use of a piece of research: generalization is an act, which is completed when someone can make sense of situations or processes or other phenomena with the help of the interpretations, which emanate from research texts. A lot of educational research describes and interprets processes, which emerge in situations and human actions, but only as a potential. You cannot claim that they always emerge in these situations or when these persons act. This creates special challenges for generalization claims. We can compare the use of a substantial portion of qualitative research with the development of a diagnostic repertoire: interpretational tools for identifying patterns in the everyday world and making better sense of the world around us. We can compare with nosography in medicine – the systematic description of deceases.

Here, I will explore the limitations of a focus on context, as a basis for generalization in more detail. There are two ways of problematizing such a generalization concept. The first argues that context similarity does not imply that the interpretation of one context must be useful in another but similar context – it might, but must not. The other argument is that interpretations of a context transcend the original context and can, and are, often useful in interpreting other contexts, which are not necessarily strikingly similar. We even use interpretations, which were originally about one kind of practice, to understand a different kind of practice. I would suggest that there is a logic here, which is quite different from the reasoning based on 'context similarity'.

According to the first argument – interpretations might, but must not be, useful in another similar context – context similarity only indicates a pragmatic potentiality, i.e. it can be practical to be alert to the possibility that an interpretation from a research study makes sense also in this new case. However, being able to generalize from a researched case to this new case is a matter of how the interpretation fits the whole case. It is the whole configuration – interpretation in context – that is the basis of generalization – an experience of a recognition of something. Here, there is no *a priori* assumption that an interpretation can be generalized to similar contexts: It is not enough. The generalization is loosely related to contexts in the sense that the researcher cannot predict in which cases the interpretation is useful, only suspect in which contexts one might look for it. It is often more a matter of 'realization' – someone who is familiar with a piece of research realizes that the original interpretation 'fits' cases they have met. The metaphor of diagnoses can again serve to illustrate the

point: The task for medical doctors is to recognize patterns that turn up during the examination and conversation with a patient, where all the pieces form a configuration, a specific diagnosis. They realize that they have a case of *x*. In one way, this line of reasoning operates with the same assumption as some of the authors in the 'context similarity' category worked with: that the task of generalization is shifted from the researcher to the audience. However, it is not about the context only but also the interpreted context, i.e. a convincing interpretation will emerge when the original text is used to interpret a new context. I operate here with the assumption that no contexts can be identified without interpreting it as something. A piece of qualitative research 'offers' a way of interpreting other cases than the original.

A difficulty for generalization through context similarity is a hidden assumption about homogeneity within a context. It is a trivial fact that different students or groups of students in the same classroom act differently. To return to the comparison with the medical doctor: simply saying that a certain context determines a specific illness is not accepted; rather, that the specific qualities in the patient's state come together with more peripheral information about circumstances and history, etc. In this case, it is also the user of knowledge who determines when a description (interpretation) is useful and applicable. Kennedy makes this point, discussing evaluation: 'The evaluator should produce and share the information, but the receivers of the information must determine whether it applies to their own situation' (1979, 672). She also points at law and clinical medicine as fields, where this is an established way of generalizing. Stake touches on this line of reasoning in a text on 'naturalistic generalizations': 'Naturalistic generalizations are conclusions arrived at through personal engagement in life's affairs or by vicarious experience so well constructed that the person feels as if it happened to themselves' (1995, 85). It seems as if Stake focuses on description with very thick descriptions – as if it were lived experience that was described. I think this is an unnecessary limitation of this line of reasoning. It is as if language, metaphors, constructions were not valid parts of humans' ways of making sense of their world (Atkinson 1990). The ideology of naturalism has without doubt been subjected to criticism (for instance, Clifford and Marcus 1986). In spite of this, Stake operates with a variant of 'generalization through recognition of patterns'. Schofield, who seems to focus on context similarity, transcends the focus on the case as representing the typical by suggesting the choice of cases, which can be the leading edge of change (1993, 103). She also discusses the choice of exemplary cases, thus giving it a normative role as an ideal (106–7). It is obvious from her suggestion that generalization here is focused on what the audience can learn from it, but it is not generalization through recognition of patterns in the empirical world. Wolcott might be closer when he gives the reader the role of 'completing' the researcher's contribution: 'The art of descriptive research, I believe, is in portraying the case at hand so well that readers themselves make the generalizations for us. They fill in or complete the pattern work that we outline only faintly' (1994, 113). Generalization by recognizing a pattern can happen even if the context-to-be-understood is different from the original study, as long as the pattern is recognizable – a somewhat odd consequence of this line of reasoning.

Critical comments

The strength of this case is that it can deal with research, where there is a loose relation between the context and the phenomenon in focus. I think this line of reasoning is more

realistic than context similarity in many cases, not least when the objects of study are processes. Giving the user 'responsibility' for generalizations highlights the issue of authority: the researcher loses authority or at least control and the audience becomes the judge of the meaningfulness of a piece of research. The audience can be both other academics and practitioners – both with varying degree of knowledge and experience of the studied phenomena. Users are not necessarily very sophisticated, which risks corrupting the original qualities of a study when it is generalized and used. The lack of a clear claim on which contexts to generalize might reduce the possibility for researchers to act as experts when facing a sceptical audience. This might be a blessing, when it reduces the halo effect of academic authority, i.e. when scholars judge in cases they do not know more about than anyone else. On the other hand – having carried out a study is supposed to generate an expertise and the respect for this – it is on the whole a fundamental argument for doing research at all. The researcher's skill in persuading the audience (as well fellow researchers and the general public) becomes a cornerstone of a generalization of an interpretation. An effective rhetoric and the precise use of language become necessary. This is nothing new – effective metaphors have always been part of successful social science (Atkinson 1990) and aesthetic qualities are very much highlighted as key qualities in qualitative research (Denzin 2000; Richardson 2000). However, this invites difficult questions about the boundaries between warranted claims and seductive language, which will not wane since there seems to be no convincing demarcation line (Phillips 1987). The dialectic in a debate about the validity of an interpretational pattern might help to expose the usefulness of it. Critical examination in such a deliberation might enhance the precision in generalization.

Final remarks

It has not been my aim in this text to reach a conclusion about a single best line of reasoning. Rather, I wanted to make the point that there are several different lines of reasoning that can be possible to use. The usefulness of each resides in the specific circumstances and purposes of each piece of research. Qualitative researchers therefore need a repertoire of possible lines of reasoning. Researchers have to find out which line of reasoning makes sense in the specific study they are conducting. Pluralism is underscored by my understanding that in some cases several lines of reasoning on generalization can be applicable to different parts of the same work. Another aspect is also the need to elaborate the reasons, when claiming that generalization is not necessary. However, nothing is perfect: I have tried to point to problems related to each of the five lines of reasoning. The Greek word 'phronesis' expresses something I think is important to consider as a final remark. The notion is about how to judge individual cases, taking into account all relevant aspects of the case. It is different to applying universal rules. A text like mine, discussing something in principle, can invite readers to draw overstated conclusions. The need for a plurality, a repertoire of lines of reasoning, is hopefully supported by the arguments presented here. However, there is no reason to believe that there are not more ways of thinking about generalizations. Campbell and Stanley write that generalization in a strictly logical sense is not possible. One has to guess: 'guesses as to what factors lawfully interact with our treatment variables, and, by implication, guesses as to what can be disregarded' (1963, 187). Any study will face the fact that it was performed in the past when it is published: It seems to be difficult to argue in a strictly logical sense that nothing has changed. Generalization is a pragmatic matter, where perfection has no place. Case

studies and experiments share the problem that they were performed with a limited number of persons and in specific contexts. In spite of this, the academic community accepts them as ways of gaining knowledge. A long time ago, Cronbach criticized the possibility of drawing conclusions from experiments in teaching: 'When we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion' (1975, 125). These examples show the general fragility in conclusions about human social life. On a more fundamental level, the problem in social science is caused by the fact that human beings are not only nature but also culture, i.e. an object of both natural science and human science. What we are investigating are creations of the human mind, i.e. formed by interpretations that are not static but dynamic. Giddens (1990) points to a 'double hermeneutics' i.e. researchers study humans who are acting on interpretations, which are sometimes produced by researchers. In the final analysis, every researcher as well as every reader must strive for wise estimates of or sophisticated discussions on the limits of the use of a specific study. Rules are collective wisdom in universalistic form, but they must be subordinated to clever judgements about the specific case. In the case of making generalizations, these wise judgements about how to deal with the specific case seem to be in great demand.

References

Ahlberg, A.W., N. Lundqvist, and G. Sörbom. 1964. *Latinsk–svensk ordbok* [Latin–Swedish lexicon]. Stockholm: Bonniers. (in Swedish)

Atkinson, P. 1990. *The ethnographic imagination: Textual constructions of reality.* London: Routledge.

Brown, A.H., R.M. Cervero, and J. Johnson-Bailey. 2000. Making the invisible visible: Race, gender and teaching in adult education. *Adult Education Quarterly* 50, no. 4: 273–88.

Campbell, D., and J. Stanley. 1963. Experimental and quasi-experimental designs for research on teaching. In *Handbook of research on teaching,* ed. N. Gage, 171–246. Chicago, IL: Rand McNally.

Chase, S.E. 2005. Narrative inquiry: Multiple lenses, approaches, voices. In *Handbook of qualitative research,* ed. N.K. Denzin, and Y.S. Lincoln, 3rd ed, 651–79. Thousand Oaks, CA: Sage.

Clifford, J., and G. Marcus, eds. 1986. *Writing culture: The poetics and politics of ethnography.* Berkeley, CA: University of California Press.

Cronbach, L.J. 1975. Beyond the two disciplines of scientific psychology. *American Psychologist* 30, no. 2: 16–27.

Denzin, N.K. 2000. Aesthetics and the practice of qualitative inquiry. *Qualitative Inquiry* 6, no. 2: 256–65.

Denzin, N.K., and Y.S. Lincoln, eds. 1994. *Handbook of qualitative research.* Thousand Oaks, CA: Sage.

Durkheim, E. 1985. *The evolution of educational thought. Lectures on the formation and development of secondary education in France.* London: Routledge & Kegan Paul.

Gamble, J. 2001. Modeling the invisible: The pedagogy of craft apprenticeship. *Studies in Continuing Education* 23, no. 2: 185–200.

Geertz, C. 1999. Thick description: Toward an interpretive theory of culture. In *Qualitative research,* ed. A. Bryman and R.G. Burgess, vol. 3, 346–68. London: Sage.

Giddens, A. 1990. *Consequences of modernity.* Cambridge, MA: Polity Press.

Glaser, B.G., and A.L. Strauss. 1967. *The discovery of grounded theory: Strategies for qualitative research.* New York: Aldine.

Greenwood, D.J., and M. Levin. 2005. Reform of the social sciences, and of universities through action research. In *Handbook of qualitative research,* ed. N.K. Denzin, and Y.S. Lincoln, 3rd ed., 43–64. Thousand Oaks, CA: Sage.

Guba, E.G., and Y.S. Lincoln. 2005. Paradigmatic controversies, contradictions, and emerging confluences. In *Handbook of qualitative research,* ed. N.K. Denzin, and Y.S. Lincoln, 3rd ed, 191–215. Thousand Oaks, CA: Sage.

Hammersley, M. 1992. *What's wrong with ethnography? Methodological explorations.* London: Routledge.

Kennedy, M.M. 1979. Generalising from single case studies. *Evaluation Review* 3, no. 4: 661–78.

Larsson, S. 2005. Om kvalitet i kvalitativ forskning [On quality in qualitative research]. *Nordisk Pedagogik* 25, no. 1: 16–35.

Lincoln, Y.S., and E. Guba. 1999. Establishing trustworthiness. In *Qualitative research,* ed. A. Bryman and R.G. Burgess, vol. 3, 397–444. London: Sage.

Marton, F. 1994. Phenomenograhy. In *The international encyclopedia of education,* ed. T. Husén and T.N. Postlethwait, 2nd ed, 4424–9. Oxford: Pergamon Press.

Mead, M. 1963. *Sex and temperament in three primitive societies,* 2nd ed. New York: Morrow Quill.

Phillips, D.C. 1987. Validity in qualitative research. Why the worry about warrant will not wane. *Education and Urban Society* 20, no. 1: 9–24.

Phoenix, A. 2004. Using informal pedagogy to oppress themselves and each other. *Nordisk Pedagogik* 24: 19–38.

Richardson, L. 2000. Evaluating ethnography. *Qualitative Inquiry* 6, no. 2: 253–55.

Scheurich, J.J. 1997. *Research method in the postmodern.* London: Falmer Press.

Schofield, J.W. 1993. Increasing the generalizability of qualitative research. In *Educational research: Current issues,* ed. M. Hammersley. London: Paul Chapman.

Shulman, L.S. 1997. Disciplines of inquiry in education: A new overview. In *Complementary methods for research in education,* ed. R.M. Jaeger, 2nd ed, 3–29. Washington, DC: American Educational Research Association.

Snyder, B.R. 1971. *The hidden curriculum.* New York: Alfred A. Knopf.

Stake, R.E. 1995. *The art of case study research.* Thousand Oaks, CA: Sage.

Strauss, A.L., and J. Corbin. 1990. *Basics of qualitative research.* Thousand Oaks, CA: Sage.

Söndergaard, D.M. 2002. Post-structural approaches to empirical analysis. *Qualitative Studies in Education* 15, no. 2: 187–204.

Thorne, B. 1993. *Gender play: Girls and boys in school.* Buckingham: Open University Press.

Usher, R. 1996. A critique of the neglected epistemological assumptions of educational research. In *Understanding educational research,* ed. D. Scott, and R. Usher, 9–32. London: Routledge.

Vogt, W.P. 1999. *Dictionary of statistics and methodology,* 2nd ed. Thousand Oaks, CA: Sage.

von Wright, G.H. 1971. *Explanation and understanding.* Ithaca, NY: Cornell University Press.

Wittgenstein, L. 1958. *Philosophical investigations.* Oxford: Basil Blackwell.

Wolcott, H.F. 1994. *Transforming qualitative data: Description, analysis, and interpretation.* Thousand Oaks, CA: Sage.