What Good are Statistics that Don't Generalize?

by David Williamson Shaffer and Ronald C. Serlin, *University of Wisconsin, Madison*

Quantitative and qualitative inquiry are sometimes portrayed as distinct and incompatible paradigms for research in education. Approaches to combining qualitative and quantitative research typically "integrate" the two methods by letting them co-exist independently within a single research study. Here we describe intra-sample statistical analysis (ISSA) as a general technique for using quantitative tools to support qualitative inquiry so as to simultaneously provide warrants from qualitative and quantitative traditions. In certain circumstances ISSA makes it possible to relax the requirement that individual participants be treated as the unit of analysis in statistical models, and thus provides justification for coding qualitative observations and drawing statistically based conclusions about observations in a qualitative context. We developed ISSA and describe it here both because it can be used as a tool for qualitative research, and because it illuminates the relationship between method and interpretation in the research traditions that it bridges. In this article, we (a) summarize key features of qualitative and quantitative research relevant to ISSA; (b) describe ISSA as an analytical technique; (c) discuss the quantitative and qualitative justification for ISSA and the nature of the conclusions that can be drawn based on it; and (d) explore the more general implications of ISSA for qualitative and quantitative inquiry.

Oh, East is East, and West is West, and never the twain shall meet,
Till Earth and Sky stand presently at God's great Judgment Seat

—Rudyard Kipling, The Ballad of East and West

Though written about a border skirmish in the "Great Game" of colonial power at the end of the 19th century, these lines from Kipling's "The Ballad of East and West" aptly describe the current schism in education research methodologies. Some have declared the end of the "paradigm wars" between qualitative and quantitative methods (Tashakkori & Teddlie, 1998, p. 1), but debates about the comparative validity and utility of research traditions[1] continue, not least in recent critical reports about the nature of education research (e.g., National Research Council, 2002). The persistence of methodological divisions is hardly surprising. Although *qualitative research* and *quantitative research* are really collections of loosely related methods and associated techniques rather than natural categories, substantive differences do exist between them. These differences result, ultimately, from differing perspectives on the nature of fundamental and problematic concepts such as "truth" and "understanding," on the kinds of claims that are worth making, and on the kinds of warrants required in scholarly inquiry.

The lines that follow Kipling's oft-quoted description of worldviews in conflict are less well known, but are also applicable to the debates about methodologies in education research. The poet reminds us: "But there is neither East nor West, Border, nor Breed, nor Birth/When two strong men stand face to face, tho' they come from the ends of the earth!" Although we prefer *thoughtful people* to *strong men*, it is in this spirit of collegiality that we write here, as a statistician on one hand, and a practitioner of qualitative inquiry on the other, both interested in understanding more deeply—and forging linkages between—the different perspectives that define our two research traditions.[2]

In recent years, researchers have explored ways to combine qualitative and quantitative research methods in education. In most cases, these efforts have amounted to declaring a "cease-fire" in the paradigm wars: letting the two methods co-exist in a single research endeavor, such as collecting survey data in parallel with focus groups; or operating in separate but mutually-reinforcing ways, such as explicitly using qualitative studies to define the parameters of quantitative investigations (Creswell, 2002; Tashakkori & Teddlie, 2003). Such pragmatic mixed-method and mixed-model studies are useful both as empirical research, and as demonstrations of the wisdom of respecting and using the strengths of different methods. However, some scholars suggest that we need to go further to develop a new post-schism paradigm for research (National Research Council, Pellegrino, Chudowsky, & Glaser, 2001). The technique of *verbal analysis* is particularly notable as an effort to integrate qualitative and quantitative methods in a more synergistic fashion (Chi, 1997).

In what follows, we describe *intra-sample statistical analysis* (ISSA) as a technique for using quantitative tools to support scientific inquiry: a technique that combines methods and assumptions of qualitative and quantitative research within a single analytical process. Briefly, in certain circumstances ISSA relaxes the requirement that individual participants be treated as the unit of analysis in statistical models. In so doing, ISSA provides a justification for coding qualitative observations and applying statistical analyses to draw conclusions about patterns of activity

This work was supported in part by a Spencer Foundation/National Academy of Education Postdoctoral Fellowship, a grant from the Wisconsin Alumni Research Foundation, and by a National Science Foundation Faculty Early Career Development Award (REC-0347000). Any opinions, findings, or conclusions expressed in this article are our own and do not necessarily reflect the views of the funding agencies or cooperating institutions.

Educational Researcher, Vol. 33, No. 9, pp. 14–25

for data sets that do not readily lend themselves to analysis across individual participants. The results of such analyses do not "generalize" in the traditional quantitative sense, but they do justify drawing statistically based conclusions about observations in a qualitative context. Conclusions drawn from ISSA are thus defensible both statistically and phenomenologically: they simultaneously provide warrants from both qualitative and quantitative traditions. In so doing, ISSA analyses can provide additional justification for the qualitative work they support.

In presenting ISSA, we want to be clear that we are not claiming that it is in any sense *more* valid than other qualitative (or quantitative) techniques. ISSA is *not* a way to make qualitative research *appear* to be more "scientific." We claim only that ISSA is another useful tool in the kit of research techniques.

We developed ISSA and describe it here partly because it is useful in the conduct of qualitative research. But we also discuss ISSA to illuminate claims and assumptions in the research traditions that it bridges. ISSA provides an occasion to clarify a fundamental premise of quantitative inquiry, separating requirements embedded in the statistical techniques from the typical interpretation of those requirements in the conduct of education research. Similarly, ISSA provides an opportunity to look at the claims of qualitative inquiry, delineating more clearly the assumptions that underlie qualitative analyses and the implications drawn from them. In bridging the paradigms of quantitative and qualitative research, ISSA thus sheds light on the relationship between method and interpretation in both traditions.

We begin by summarizing the key features of qualitative and quantitative research relevant to ISSA, with a brief discussion of existing techniques for mixed qualitative-quantitative analyses. Next, we describe ISSA as an analytical technique, explaining its quantitative and qualitative justification and the nature of the conclusions that can be drawn from it. We ground this description in a hypothetical example of an ISSA analysis based on empirical work. Finally, we discuss the implications of ISSA for our understanding of the nature of qualitative and quantitative inquiry.

Qualitative and Quantitative Paradigms

There is not space—nor would we be so bold as to try—to provide a comprehensive summary of quantitative and qualitative research theories and techniques. What follows is an overview of the assumptions of quantitative and qualitative analyses upon which ISSA operates, and a brief discussion of the relationship of ISSA to other techniques that combine these two traditions.

There are many techniques for quantitative analysis of data, including both descriptive statistics and a variety of ways to generalize from observed data. Methods for making generalized claims from a set of observations—including common techniques such as correlation, regression, and analysis of variance—all depend on the idea of *sampling*. The assumption underlying sampling is that the results we have observed (the sample) are drawn in an unbiased way from some larger population. The statistical question is whether the characteristics of the sample reflect characteristics that hold in general for the larger population from which our sample was taken.[3] In the context of education research, the sample is often a collection of individual students about whom we collect data.[4] For purposes of statistical analysis, we view those students as coming from an idealized population of students "similar to those in the study"—which is why quantitative researchers are so concerned about recruitment, self-selection bias, and other issues that limit the kind of students we can claim are "similar" to those about whom data is collected.

Quantitative analyses warrant claims that observations made for a specific set of individuals generalize to "all students like these" by distinguishing between "true effects" (or relationships) in the population as a whole and the normal (and approximately normally distributed) random variations among individuals.[5] The goal of the analysis is to determine whether the effects seen in the sample are likely to reflect "true effects" for "students in general" and not merely chance aggregations of random characteristics of the individuals selected for a particular sample. Such claims depend on looking at individual students as a random-effects unit drawn from a larger population of similar individuals. That is, the individuals in the sample must be the unit of analysis for statistical purposes. The impact of systematic effects on statistical tests increases in larger samples because such effects follow a consistent pattern (which is why they are "systematic"). The effects of random errors on statistical measures, in contrast, tend to cancel out as samples increase in size. A larger sample (studying more individuals) thus provides more *power* to a quantitative analysis because larger samples make it possible to warrant claims about more subtle effects and relationships in the data.[6]

Such non-contextualized generalization is not the goal of qualitative research. Qualitative researchers typically reject the notion that there exist "true effects" that can be attributed separately to subjects, actions, or interventions; rather, they take as a fundamental premise the idea that observations are produced through the contingent interaction of participant and researcher.[7] There are a number of techniques and traditions in qualitative research, but the overarching goal of qualitative inquiry is to provide some form of what Geertz (1973b) popularized as "thick description."[8] A thick description of a context is an attempt to understand how and why events unfolded in a *particular* place and time, from which the researcher can draw inferences about *specific* participants' experiences, assumptions, emotions, and understandings *in a given setting*—and thus why they acted in the ways they did. In other words, qualitative inquiry is useful for understanding causal connections in the lived experience of participants, and the inferences from qualitative analyses are typically used to provide a framework for more subtle and sophisticated interpretation (or reinterpretation) of data in other contexts. In this way, qualitative studies build upon one another, providing increasingly nuanced understanding of phenomena.

Some theorists argue that descriptions of causal mechanism are naively realist and therefore suspect from a qualitative perspective (Tashakkori & Teddlie, 1998); however, any qualitative analysis has to assert some claim to being more a portrait of the experience of participants than a reflection of the biases of the researchers. Qualitative researchers use a variety of techniques to deal with bias—not to "control" or eliminate it, but rather to understand and document its effect on the account presented. Methods to warrant a qualitative account as a representation of the experience of participants include triangulation from other sources and accounts, presentation of findings to participants for feedback, and, most commonly, demonstrating the clarity and consistency of phenomena described in data collected. One particularly useful technique for warranting such claims is to gather data until additional observations confirm existing hypotheses rather than lead to new insights—a condition referred to as *theoretical saturation* (Glaser & Strauss, 1967; Strauss & Corbin, 1998). To that end, a key source of power in a qualitative analysis comes from collecting a great deal of information about a small number of subjects, trying to understand events and persons in depth through a large corpus of observations.

Both quantitative and qualitative techniques thus gain increasing power by collecting more data. However, the concept of *more powerful analysis* differs in the two traditions. Indeed, the concepts of analytic power in qualitative and quantitative research as traditionally understood are contradictory. Qualitative analysis is more powerful when it presents a richer portrait of persons and their actions in a particular context. Quantitative analysis is more powerful when it can make claims to generality about more subtle differences in observed data. Quantitative data are typically not powerful in a qualitative analysis because they are too thin: there is not enough information about the individual participants to warrant thickly descriptive inferences. Qualitative data are typically not powerful in a quantitative analysis because the number of individuals studied—the size of the sample—is typically small. In the world of finite resources where all research takes place, the impetus to collect data on more individuals in a quantitative study inherently conflicts with the need to collect large amounts of information about each participant in a qualitative study.

In recent years, researchers have begun using both quantitative and qualitative techniques, recognizing that different methods of analysis are useful for addressing different kinds of questions. Guides to mixed-method research describe a taxonomy of studies using quantitative and qualitative methods. Tashakkori and Teddlie (1998) look at studies in which quantitative and qualitative techniques are used sequentially or in parallel, with equal or differential status in addressing research questions, in the same phase or in different phases of a single study. They describe how quantitative analysis might identify subjects for a qualitative study; how qualitative interviews might provide additional insight into processes identified through quantitative analysis; how qualitative analysis might generate hypotheses for quantitative study; and how quantitative and qualitative data might be collected simultaneously. Whether the methods are used concurrently or sequentially, however, the interaction is between methods; the methods remain distinct. (See also Creswell, 2002; Tashakkori & Teddlie, 2003.)

Chi (1997) describes a similar taxonomy, and argues that *verbal analysis* provides a more integrated approach to combining quantitative and qualitative methods. Verbal analysis involves segmenting and coding verbal protocol data, and then using a variety of techniques (including quantitative analysis) to identify and interpret patterns in the coded data. Chi's explication of her method provides an exemplary guide to the process of quantifying verbal protocol data, including such issues as reduction or sampling of protocols; segmentation of data; development of a coding scheme; operationalization of coding; depiction of the coded data; analysis of the coded data; and interpretation of results.

In combining qualitative and quantitative methods, the primary methodological concerns Chi addresses are about the translation of qualitative data into quantitative form: segmentation, coding, and procedures for maintaining the consistency needed to claim that observations are independent and equivalent for the purposes of statistical analysis. To that end she provides a thorough explanation of her methods, including numerous "recommendations, technical details, and caveats" (1997, p. 303). However, Chi's discussion does not address in depth how and why one might use statistical analyses to support qualitative inference—nor the questions that such an integration raises about the nature of qualitative and quantitative inquiry.

ISSA, then, extends the concept of verbal analysis to encompass qualitative data of any kind, including observations as well as interviews, task analyses as well as think-aloud protocols, video and field notes as well as audio transcripts. More importantly, ISSA provides a theoretical justification for the use of statistical analyses to support qualitative inference—and thus an occasion to reexamine the assumptions of quantitative and qualitative research traditions themselves.

Intra-sample Statistical Analysis

In this section, we describe a hypothetical study to frame the key issues, processes, and assumptions of ISSA. MathStudio (our hypothetical study) is a composite of several projects (Shaffer, 1997, 2002, 2003, 2004), which we have combined for rhetorical simplicity and altered somewhat for conceptual clarity.

Example, Part 1: Learning Math through Design

MathStudio was a 13-day after-school program in which 12 middle school students used computer software to create mathematical designs for display in a museum exhibit. Each day, while students were working on projects, three mentors conducted clinical interviews with the students. These interview data were combined with records of student work to create a *work history* for each student: a record of what he or she did in MathStudio, annotated with the student's interpretations and mentors' observations.

These work histories could be analyzed qualitatively for patterns in the relationships between these students and their mentors. For example, let us assume that in looking at the data the researchers saw how, during extended design activities, those relationships came to incorporate desires for self-expression, feelings of frustration, and ritualized forms of critique, to turn interactions between students' intent and the obstacles they encountered into opportunities for mathematical development. That is, let us assume the researchers developed a *grounded theory* (Glaser & Strauss, 1967; Strauss & Corbin, 1998) about the processes through which these students in MathStudio developed mathematical insights. This interpretation would be based on patterns of activity identified in the observations of student work: on the complex system of action and interaction among students, mentors, tool, activity, and domain that seemed to make it more likely that students would say that they had learned something about mathematics. For the sake of argument, let us assume that the pattern was:

1. A student started out trying to "say something" about mathematics with his or her design—that is, a student started out with an *expressive intent.*
2. The student became frustrated when he or she ran into problems making the design "say" what he or she wanted it to—that is, the student encountered *expressive obstacles*.

3. Motivated by the original expressive intent to overcome the expressive obstacle, the student brought those problems to a mentor for help, and the mentor used a pattern of questioning common in the practice of graphic design—that is, the student and mentor enacted a *participant framework of design*.
4. As a result of examining the expressive obstacle in the context of the participant framework of design, the student gained insight into the mathematical principles at work in his or her design—that is, the student developed some kind of *explicit mathematical understanding*.

This grounded theory could be explored in a number of ways using qualitative techniques. The researchers could describe a single case or multiple cases in depth, choosing examples for which data on student and mentor thinking were particularly rich. The researchers could describe a set of contrasting cases in which the observed progression broke down at different points. The researchers could describe this progression and ask students and mentors to give feedback on how accurately it reflects their experience in the program. The researchers could count the number of students who went through the progression described at some point in the workshop, showing that some or all of the students went through the progression (perhaps more than once) during the project. As Chi (1997) describes, the researchers could segment the original data into *design episodes*, establishing criteria for when students started and stopped working on a particular project. These segmented episodes could be coded, and for each student researchers could count: the number of design episodes, the number of episodes in which the progression occurred, and the number of episodes in which the progression led to mathematical insight. They could then produce descriptive statistics on the frequency and effect of the pattern observed.

Any of these techniques could be used to support the original claim, and ideally the researchers would use more than one means to demonstrate that the progression found in the data reflects the experience of the participants in MathStudio rather than some original (perhaps implicit) assumptions of the research team.

There are also quantitative analyses that the researchers could perform to warrant claims about the pattern observed. Having segmented the data into design episodes, the researchers could code for occurrences of elements of the pattern described: *expressive intent*, *expressive obstacles*, *participant framework of design*, and *explicit mathematical understanding* as described in the grounded theory.

With 13 days of work and perhaps an average of 5 design episodes per day for each participant, the data might consist of 780 design episodes—a number of observations large enough to leverage the power of statistical techniques. However, these 780 observations would be repeated observations of the same 12 individuals, so from the point of view of typical quantitative analyses, the sample size would be only 12. To use statistical techniques in the framework of typical quantitative analyses, the researchers would likely aggregate the data, computing the frequency of events for each participant, and test to see whether there were statistically significant correlations in the frequencies of these codes across individuals.[9]

This presents two problems. Collapsing the data from 780 episodes into frequencies of events across 12 individuals substantially diminishes the power of the analysis. It is true that conclusions from the quantitative analysis of these 12 students could be generalized to a statistical claim that "all students like those 12" would act in similar ways in the same experiment. But only the broadest patterns would be likely to achieve statistical significance with such a small sample—and it is quite conceivable that there would be no statistically significant effects even if the pattern could be supported by other qualitative techniques.

More problematic, though, analyzing the data in this way would actually change the nature of the research question. The original grounded theory was a claim about the behavior of these 12 participants—a thick description of what they did and why—rather than a general claim about how students behave during mathematical design activities. The original claim was about patterns of activity and motivation within each design episode—and perhaps unfolding across design episodes—rather than a "true effect" attributed to the individuals or the situation.

What our hypothetical researchers want from a quantitative technique, then, is not a way to generalize to a larger population of students, but a means to provide additional warrants for their claim about the processes identified through qualitative analysis. ISSA is a tool for providing such quantitative support for qualitative claims.

ISSA, Part 1: Relaxing a Key Assumption

The qualitative claim in MathStudio is about patterns of activity within design episodes; thus, our researchers might be tempted to conduct statistical analyses with episodes as the unit of analysis. To do that, they would need to account statistically for the fact that the 780 episodes were the work of 12 distinct students, each with particular interests, inclinations, skills and propensities. One possibility would be to analyze the design histories for each student separately; this would provide information about the patterns of activity for each student individually. Alternatively, the researchers could treat students as "fixed effects" in the analysis by including dummy variables for each student in a standard regression analysis.[10] Doing so would allow the researchers to model the patterns of activity in this group of students as a whole. However, the rules of quantitative analysis, as typically interpreted in educational settings, would say that such analyses would be inappropriate because the individual subjects are not the unit of analysis and thus the results will not generalize.

The problem is the assumption about the population from which the sample is drawn. A traditional quantitative analysis requires that the population be a hypothetical collection of "all students like the ones in the study." Instead of a population of "students like these", ISSA posits a hypothetical population of "observations about what these particular students did in this particular situation." The observations we have recorded are thus a sample from all the things that we might have recorded about these students in the given context from a particular perspective: an ideal population of observations. Positing an ideal population of observations that could have been made about particular participants in a particular context does not imply there is a set of things that "really happened"—only that there is a larger set of things that might have been recorded of which we have a subset.

Of all the observations that could be made about a group of specific people in a particular context, a researcher records a subset and wants to make a claim about patterns of activity reflected in those observations. If this were not the case, there would be nothing to analyze qualitatively or quantitatively. The ideal population of observations represents the full set of possible observations (by a given observer or observers under specific conditions), recording from a particular perspective the activities in which a particular set of participants could be seen to engage in a given context. ISSA posits that the particular observations in any study are a sample from this larger pool of possible observations, and a statistical analysis with the observations (rather than individual students) as the random unit of analysis supports an inference that the specific pattern of behavior observed reflects the larger patterns of activity that could be observed in these particular students in the given context.

In taking such a step, ISSA sacrifices the ability to generalize beyond this group of students on this particular occasion. In exchange, however, ISSA makes it possible to use statistical techniques to gain insight into patterns of activity, and to support inferential claims based on qualitative methods in situations that do not otherwise lend themselves to investigation with individual participants as a unit of analysis. In this sense, ISSA does not supplant other quantitative approaches to qualitative data, such as verbal analysis. Rather, ISSA makes it possible to derive, for generalization within a particular group of students, more statistical power from studies with large numbers of observations for a small number of subjects. The results will not generalize to a larger population of students, but they will provide additional warrants for qualitative claims about smaller effects than can be analyzed using traditional quantitative analyses designed to generalize to a larger population of individuals. Practically speaking, this means that we can generate a qualitative inference, code the data appropriately, and use traditional statistical techniques to analyze data within our sample, rather than between individuals.

Example, Part 2: What Can We Conclude?

In MathStudio, ISSA thus allows our researchers to construct a statistical model of the grounded theory. They might, for example, use logistic regression analysis, taking the individual design episodes as the unit of analysis. The presence or absence of *explicit mathematical understanding* might be the outcome variable; the presence or absence of *expressive intent*, *expressive obstacles*, and enactment of a *participant framework of design* might be the predictors; and individual students would be treated statistically as fixed effects.

The results of such a regression analysis would be a model that would give odds ratios related to each predictor, describing the increased (or decreased) likelihood that an episode would lead to insight based on the presence of and interactions among elements of the pattern identified through qualitative analysis. For example, the model might show that students were 4.6 times more likely to develop explicit mathematical understanding in a design episode when they engaged in a participant framework of design, and 2.2 times more likely to develop explicit mathematical understanding when they encountered expressive obstacles.[11] The regression would also, of course, give the statistical significance of those odds ratios. The model could also include terms to represent the interaction of elements in the grounded theory. The significance of those interaction effects would support a claim that for these students, the development of mathematical understanding in MathStudio depended on the relationships among expressive intent, expressive obstacles, and the participant frameworks of design.

Any number of other statistical techniques could also be used with the episodes as the unit of analysis to support observations made in the original qualitative data, producing estimates of effect size and significance. The researchers could thus test statistically whether the pattern observed is a property not just of *these particular observations*, but of *any set of observations* of these students made under the same conditions. They would not be able to make any claims about whether the pattern observed would hold true for students other than those who participated in the program—but, as above, that is not the claim the original qualitative analysis was trying to make.

ISSA, Part 2: Exchangeability

In this hypothetical logistic regression model, the researchers treated the individual students as fixed effects to help account for the fact that the 780 episodes were not independent observations but rather repeated measures of the same 12 students. A fixed-effects model removes the variability due to repeated sampling of individuals. It deals with the intercorrelation among observations of each individual—that is, the fact that each individual may have his or her own particular strengths, weaknesses, interests, or proclivities that systematically affect his or her behavior. However, to control for the potential rise in Type I errors (false positives or inflated estimates of significance) due to repeated sampling, the researchers have to make a further assumption. ISSA lets the researchers focus on the episodes as if they were random units of analysis by substituting the more general assumption of *exchangeability* in place of the criterion of *independence* typically applied in discussions of quantitative analysis (Kingman, 1978; Lindley & Novick, 1981; Mislevy, 1996).

Briefly, exchangeability means that the units of an analysis (the observations or design episodes in our hypothetical example) are functionally independent for purposes of the analysis. While this is relatively easy to describe in principle, it takes considerable work, care, and caution to establish in practice. The researchers use statistical techniques to control their ISSA analysis for the within-subjects effects of individual students—and, depending on the nature of the pattern observed, perhaps other potential intervening variables such as what time of day the episode occurred or whether the episode was the first, middle, or last design episode in an extended investigation. The researchers then posit that the episodes (the units of analysis) are interchangeable: after accounting for all of the reasonable factors that might create intercorrelations among episodes there is no reason to believe, a priori, that episodes differ in some systematic fashion other than in ways described in the qualitative analysis and represented by the variables that quantify the grounded theory being examined.[12]

In a sense, this approach is not substantively different from a more traditional quantitative analysis, which might assert that after controlling for race, gender, and socio-economic status, individual students are independent units of analysis for the purposes of evaluating an educational intervention.[13] Technically speaking, independence is actually a special case of exchangeability in which certain probabilities are easier to compute (Draper, Hodges, Mallows, & Pregibon, 1993; Lindley & Novick, 1981).[14]

In a typical quantitative analysis, choices about how to use variables to describe the data, and what intervening variables might need to be considered, are ideally based on prior theoretical understanding of the situation in question. In ISSA, the qualitative foundation for the hypotheses being examined provides phenomenological support for those decisions.

Example, Part 3: ISSA and the Virtue of Disaggregation

For our hypothetical researchers ISSA thus potentially provides support for a wider variety of qualitative claims than traditional quantitative analyses. In so doing, ISSA makes it possible for researchers to frame questions that are more closely aligned with the qualitative claims they are investigating, and thus potentially avoid statistical anomalies such as Simpson's paradox (Lindley & Novick, 1981) that can produce ambiguous or even contradictory results when data are summarized by individual students for the purposes of analysis.

For example, in MathStudio let us assume that students were given a pre and post test that included four word problems based on complex spatial situations—and that students completed these problems using *think aloud* protocols (Chi, 1997) with two isomorphic forms of each problem randomly assigned to the pre or post test for each student. An example might be a problem that describes a child's bus route from home to school and asks students to compute the "straight line" distance from home to school. Again hypothetically, the researchers might observe a pattern in students' responses to these problems. Let us assume the grounded theory that emerges from qualitative analysis is that before the workshop, students were more likely to solve such problems if they remembered a similar problem from mathematics class in school, whereas afterwards they were more likely to solve the problems if they could relate to the problem situation personally and then use that personal knowledge to draw a realistic diagram as a basis for their solution.

The researchers could code the protocol data and warrant this pattern using typical quantitative techniques with the individual students as the unit of analysis—most likely using an analysis of variance or set of correlations. But one can easily imagine a situation in which summary statistics might obscure or distort the qualitative pattern. For example, a summary of a portion of the posttest data might look something like Figure 1. The pattern of check marks in columns 3 and 4 suggests that students' responses to individual problems support the researchers' grounded theory that in the post test students used personal experience to help them solve problems correctly. However, the summary statistics in columns 5 and 6 show a *negative* association between use of personal experience and correct problem solving: students who related more problems to personal experience did less well overall than students who related fewer problems to personal experience. Although this is admittedly a trivial example and the confound may be easy to see in this small and simple subset of the data, the same class of problems can occur in more subtle but equally serious ways in larger and more complex sets of observations (Lindley & Novick, 1981).

Just as the qualitative claim about how students learned in MathStudio was ultimately about the nature of their design episodes, so the pattern in how students solved problems is ultimately a claim about activity within each problem. ISSA allows

Student	Problem	Related problem to personal experience	Answered correctly	Related problem to personal experience	Answered correctly
A	1	✓	✓	2	2
	2	✓	✓		
	3	X	X		
	4	X	X		
B	1	X	X	2	2
	2	X	X		
	3	✓	✓		
	4	✓	✓		
C	1	X	X	3	2
	2	✓	✓		
	3	✓	✓		
	4	✓	X		
D	1	✓	✓	3	1
	2	✓	X		
	3	✓	X		
	4	X	X		

FIGURE 1. *A hypothetical subset of data on answers to post test questions shows how patterns observed can be obscured in an analysis of aggregated data.*

our hypothetical researchers to use the solutions to individual problems as a unit of analysis to warrant such claims, and even for this very small subsample of the hypothetical MathStudio data, an ISSA analysis could easily support the grounded theory. Such an analysis would *not* support a claim that there is a "true effect" for "any students like these" who participated in MathStudio. However, it would support an inference that the pattern observed is a property not just of these particular observations, but of any observations we might have made of these students in appropriately similar conditions. Put another way, ISSA predicts that further observations would produce patterns similar to those seen in the observations at hand[15]—that is, ISSA warrants claims about theoretical saturation.

ISSA, Part 3: Utility

The claims that ISSA supports are thus less broad than those warranted by a traditional quantitative analysis. However, providing such statistical support is not therefore trivial. Many inferences will not withstand a test of their ability to generalize to the ideal set of all possible observations of a particular group of participants in a particular setting. The statistical justification ISSA provides is thus a potentially useful warrant for claims based on qualitative analysis.

The observations in any qualitative study are necessarily a subset of all the things that might have been observed using a particular set of tools and techniques in a particular setting. From this subset of all possible observations, a further subset is extracted to form the basis of qualitative inferences, since no qualitative analysis accounts for all of the observational data in equal measure. The presentation of such a qualitative analysis necessarily generates a further subset of this subset of this subset of the original data—namely, the specific examples used in any report of research—since it is neither possible nor desirable to reproduce the entire dataset in the presentation of findings. (See Figure 2.)

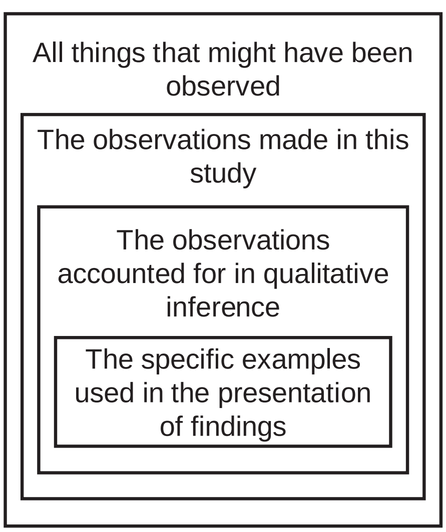

FIGURE 2. *A schematic representation of how any report of research uses a subset of all observations that might have been made in a given context from a particular perspective.*

Any thick description is thus in the end a particular story about what happened in a specific situation, and researchers need to provide warrants for their claim that the story reflects the experience of the participants more than the original biases of the research team. As a colleague in our department says: The question is, what makes this particular story better than any other story that we might tell about the same events?

By showing that grounded claims about a specific set of observations generalize to the ideal set of all possible observations of a particular group of participants in a particular setting, ISSA supports other techniques to warrant qualitative claims that an interpretation of events reflects the experience of those participants in some meaningful way. In so doing, ISSA relaxes only a single condition in the usual requirements for quantitative inquiry. Thus, all of the usual rules, techniques, and implications of quantitative analysis still apply—with, of course, the very notable exception that ISSA results do not generalize from a specific set of students to all students of a given type, but rather from a specific set of observations to all observations of a given type. That is, ISSA generalizes *within* a sample, warranting patterns inferred in the observed data, rather than *from* a sample to "true effects" in some larger population of individuals.

Discussion: The Fair Sample

All of which brings us to the question: What good are statistics that don't generalize? The utility of ISSA to support generalization within a sample of individuals suggests that the conundrum of our original question is not in the nature of statistical methods themselves, but rather in our understanding of *generalization* in education research. This, in turn, provides an opportunity to reexamine the premises of qualitative and quantitative research methods—and perhaps to find common ground that begins to bridge the intellectual chasm that appears to loom between them.

In this broader sense, ISSA revisits a landscape explored by Goodman (1978). At the core of Goodman's intellectual program was the search for an answer to the intellectual, methodological, and cultural divide between "art" and "science." A key element in Goodman's common language for artistic and scientific phenomena was the concept of the *fair sample*. Goodman used swatches from a bolt of cloth to describe this central idea. He argued that although any swatch would contain a portion of the whole pattern, some swatches would give a more accurate picture than others of the cloth's overall design. Adapting Goodman's original illustration, Figure 3 shows five swatches from a Navajo rug.[16] Each square is the same size, as in Goodman's example, and although none contains the whole pattern, swatch (c) in the upper right appears to be the fairest representation of what we might expect, based on this collection of swatches, the whole rug would look like. Based in part on such an example, Goodman defined a fair sample as "one that may be rightly projected to the pattern or mixture or other relevant feature of the whole or of further samples" (1978, p. 135).

In other words, swatch (c) is a fair sample because it provides the most faithful (under the circumstances) representation of the larger pattern, and thus gives us the most information about what we might expect to see in future samples. One way to interpret such a claim is that these swatches are observations from various perspectives of some larger image. We infer the overall pattern

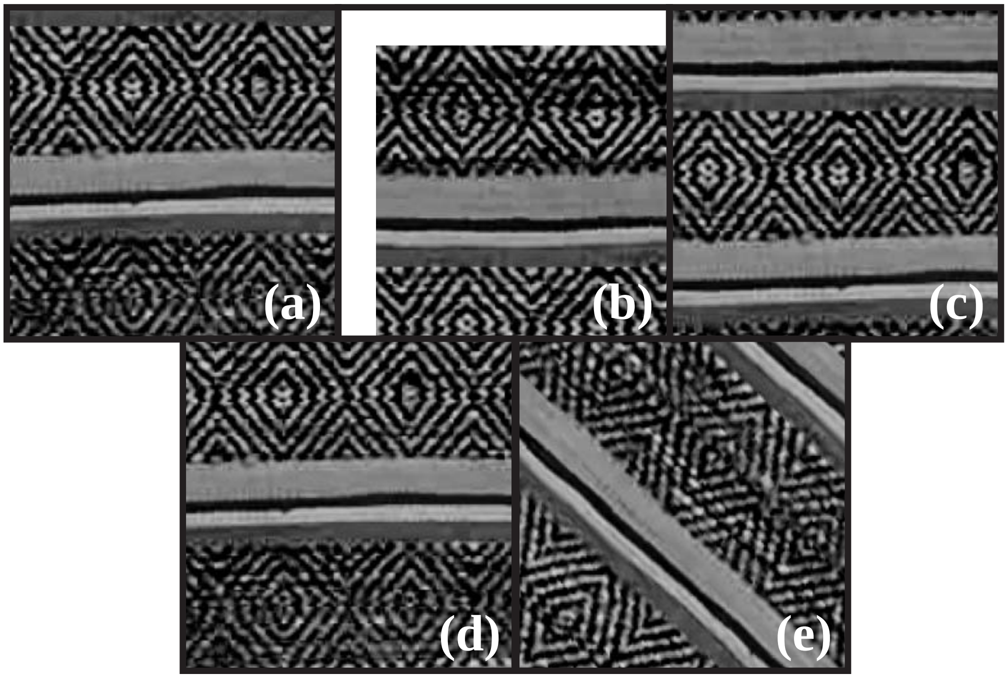

FIGURE 3. *Five swatches taken from an image of a Navajo rug recreate Goodman's example of a fair sample. Swatch (c) is a fair sample because it provides the most information about what we might expect to observe in future samples.*

from the collection of swatches—with additional swatches adding to our confidence in that inference. We conclude that swatch (c) is a fair sample in the sense that it provides the clearest representation of our inference.

The process is thus similar to a typical qualitative analysis, in which repeated observations of a small set of persons and events help us build richer portraits and more compelling claims about causal relations and lived experience. An account of qualitative inquiry is presented as a fair sample in the sense that it provides an accurate representation of these claims. Perhaps most important, as in any qualitative study, our conclusions about the pattern in this Navajo rug would be necessarily limited to inferences about the single rug in question. With a sufficient number of swatches, we might develop a rich understanding of the pattern on the rug, and perhaps even of the process through which it was created: for example, on closer examination we might see details about the fibers used, which might provide clues about how and when the rug was woven. With a rich enough knowledge about Navajo textiles, we might even discover that, based on the age, location, and techniques used in this particular rug, two groups of Navajo weavers had a greater degree of cultural exchange at an earlier time than was previously assumed.

To expand the metaphor to include a typical quantitative analysis, let us extend Goodman's example to a second collection of five swatches, this time taken from *different* Navajo rugs[17] (Figure 4).

Which swatch is now a fair sample may be less clear, because the variety of Navajo rugs writ large is greater than the variety in the pattern of any single rug. More important, though, the kind of question we would address with this new collection of swatches is different from that addressed in our first sample. This sample is a (hypothetically random) selection from some collection of Navajo rugs. With limited information about any one of the rugs, the conclusions we draw will not be detailed descriptions of any single pattern, but rather inferences about the larger population of textiles from which they were drawn. We might (rightly) conclude from this sample that traditional Navajo designs are often—though not always—based on a variety of abstract forms in symmetric patterns. We might again choose the same swatch (swatch [b] in Figure 4) as a fair sample, but it would now be a representative not of the rug from which it was taken, but of Navajo weaving in general.

The point of this extended example is not to argue that one form of inference is better or worse. Rather, these two examples illustrate the nature of qualitative and quantitative inquiry in light of ISSA. One can imagine a coding scheme that would allow us to draw statistical inferences about Navajo tapestries from a large number of samples such as those in Figure 4. We might, for example, compute for each swatch the average number of repetitions of its motifs, the number of geometric and representational motifs used, and so on. Such a claim would depend on all of the assumptions of randomness and independence of a typical quantitative sample and would represent a relatively straightforward application of quantitative analysis.

ISSA suggests that a similar process could be carried out with a set of swatches taken from the same rug (as in Goodman's original example, Figure 3). In the typical quantitative case (Figure 4), the statistical analysis is based on a sample drawn from a population of different rugs, and therefore the statistical inferences apply to all rugs from that larger population. In the typical qualitative case (Figure 3), ISSA takes the sample to be a collection of repeated observations of the same rug and allows us to draw

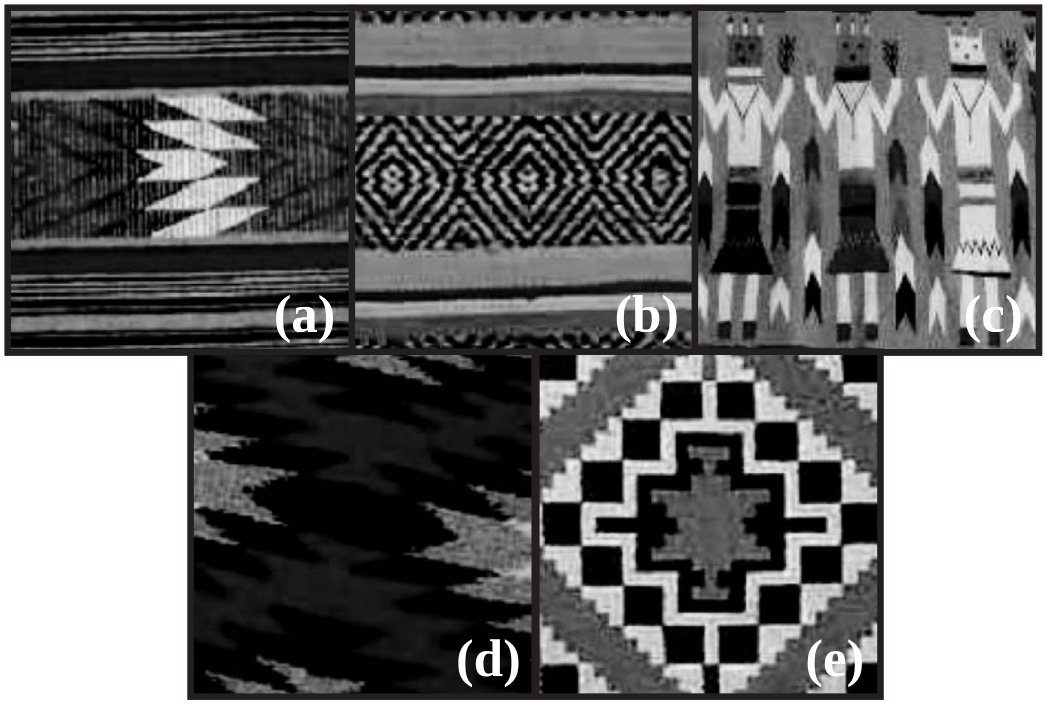

FIGURE 4. *An extension of Goodman's original example: swatches from different Navajo rugs. Swatch (b) might still be a fair sample, but of Navajo weaving in general rather than of the rug individual from which it was taken.*

statistical inferences that apply to the particular rug from which the observations were drawn.

ISSA further suggests that these two perspectives can be combined. We can imagine a small collection of rugs from which we make repeated observations—perhaps multiple swatches from different rugs (Figure 5). With appropriate statistical techniques to account for the effects of intercorrelation among observations, ISSA could help warrant qualitative inferences about this particular collection of rugs without necessarily generalizing to all Navajo rugs. If this collection of rugs came from the same location, or even the same weaver, ISSA could conceivably help support conclusions about textile production in a particular time and place by a particular person or group of people.

In warranting claims of this sort, ISSA highlights an aspect of qualitative inference that is critical regardless of the techniques used: there have to be rugs from which observations are drawn.

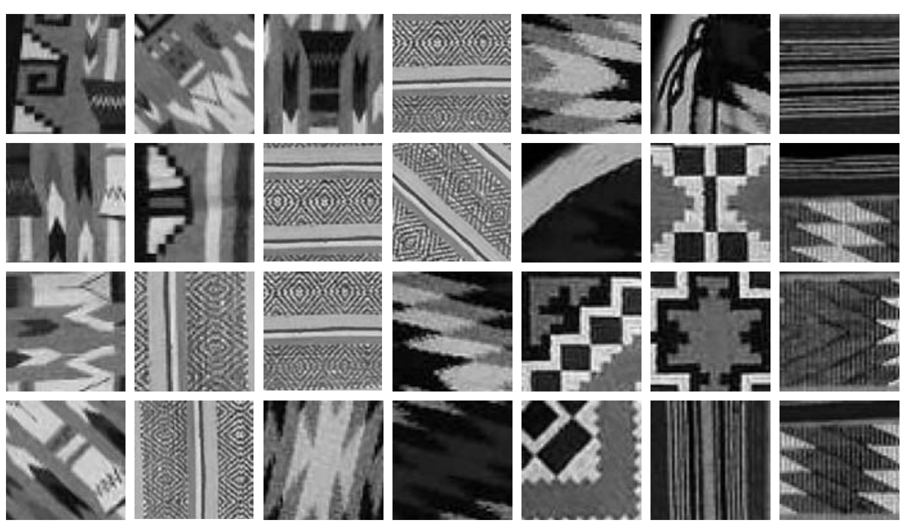

FIGURE 5. *A collection of swatches (observations) from a small collection of rugs. With appropriate statistical techniques to account for the effects of intercorrelation, ISSA can help warrant qualitative inferences about the collection without generalizing to all rugs.*

The easy caricature of qualitative inquiry's respect for the contingent nature of inferential claims is that qualitative studies imply complete relativism and are therefore unavoidably subjective. It may be true that all experience is a collection of particulars, selected by our biases, understandings, and assumptions. However, the challenge of qualitative analysis is to make sense of useful regularities within such contingent experience. This does not imply that we can make objectively true or false claims. But ISSA does highlight the sense in which any qualitative inquiry requires an assumption of consistency over time. There has to be some coherence that underlies the observations made; otherwise, there would be nothing to infer, no story to tell. Qualitative claims necessarily posit some underlying psychic and causal unity for persons and events being observed, and ISSA—like any technique for supporting qualitative claims—is an attempt to demonstrate that inferences based on those observations reflect that coherence as viewed from a particular perspective. In turn, this underlying belief in the broader psychic unity of humankind (Geertz, 1973a) is the basis for our supposition that the understanding we gain from locally observed phenomena may apply elsewhere. It is in this sense that qualitative analysis forms a basis for further investigation and action: we use understanding of one situation as a place to begin looking—or to start planning—in another situation that we believe to be similar in some meaningful way.

ISSA likewise sheds light on the assumptions that underlie quantitative inquiry. The easy caricature of quantitative inquiry's respect for experimental and statistical attempts to account for confounding influences on generalizable claims is that quantitative studies imply the search for an objective "truth" and are therefore naively positivist. ISSA suggests that there is utility in using statistical techniques beyond generalizing to large populations based on smaller groups of individuals. Inferential statistics do not require individual actors (whether students, groups, classrooms, schools, districts, or nations) as the unit of analysis—unless one is interested in making particular kinds of generalizations. ISSA shows that statistics are not inherently methods for identifying "true effects" in a larger population of individuals, but rather means to warrant claims about generalizability. ISSA reminds us that the utility of such generalization is not a statistical question, and that we therefore cannot eliminate problems of contingency by relying on experimental design alone.

In other words, ISSA suggests that qualitative and quantitative methods are both, ultimately, methods to warrant presentation of a fair sample. They are both attempts to project from a finite set of information to some larger population: a population of like individuals in the case of typical quantitative inquiry, or a collection of like observations in qualitative analysis. In showing that the same statistical techniques, properly applied, can be used to generalize in both cases, ISSA reminds us that in any kind of inquiry it is a mistake to regard technique as synonymous with interpretation—an unfortunate legacy, perhaps, of sometimes formulaic interpretations of statistical methods. Rather, the goal in any analysis is to match technique to inference, claim to warrant. The questions facing a researcher are always: What questions are worth asking in this situation? What data will shed light on those questions? And what analytical methods will warrant data-based claims about those questions? Answering these questions is a task that necessarily involves a thorough understanding of the strengths and weaknesses of a range of quantitative and qualitative techniques.

In theoretical terms, then, ISSA suggests that we should be careful neither to demonize nor to fetishize any particular methodology at the expense of the more complex task of understanding the relationship between analysis and interpretation: that is, what claims can be supported by a particular investigation. No technique—not even randomized controlled trials, which have received so much attention of late in education circles—provides a universal prescription for truth.

In practical terms, ISSA may be of limited utility in the quantitative tradition, where generalization to a larger population of individuals is the measure of utility. Qualitative inquiry, on the other hand, is not interested in generalization in this sense. As Geertz (1973b) so aptly explained: "The essential task of theory building here is not to codify abstract regularities but to make thick description possible, not to generalize across cases but to generalize within them" (p. 26). ISSA provides a justification for using statistical techniques to address this challenge.[18] In so doing, it is one tool (among many) that researchers can use to support claims that a particular pattern provides a meaningful explanation of a collection of observations made about a given set of events.

We hope that ISSA is a useful research tool, and moreover a step towards understanding—and hopefully overcoming—the ongoing schism between qualitative and quantitative inquiry. In closing this discussion of ISSA, we think it is perhaps even more important to note that whatever small progress ISSA represents came from a long series of conversations conducted in a spirit of mutual respect. The current climate of détente in research methods might be justified by Frost's ironic aphorism "Good fences make good neighbors."—, although recent assaults on all but a small set of chosen techniques might make some question whether even this spirit of benign neglect will continue. With a modicum of tolerance, qualitative and quantitative inquiry can exist side by side, and even be coordinated in the conduct of research. But we believe there is much to be gained for both methods—and for our understanding of research more broadly—from the attempt to close the intellectual distance between these two modes of inquiry. As Frost suggests: "Before I built a wall I'd ask to know/ What I was walling in or walling out." We were only able to begin to answer such question by starting the conversation with the assumption that each of our respective traditions, in the hands of thoughtful researchers, is a useful tool for investigating complex phenomena.

NOTES

[1] Yu (2003) discusses in some detail the notion of "research traditions." Here we follow common usage to frame the issue in the context of current methodological debates.

[2] Which of us is "East" and which "West" we leave as an exercise for the reader.

[3] Technically speaking, the statistical question is whether what we have observed in the sample is *not* true of the population as a whole—that is, we examine the null hypothesis. If it is *un*likely that the effects observed are *not* true of the population as a whole, we reject the null hypothesis and infer that the effects observed in the sample hold for the population as a whole. Although logically correct, the double negative is, sadly, as problematic in statistics as it is in prose.

[4] This is not universally the case, as macro-level studies focusing on the effects of curricula or educational programs often take classrooms or schools as units of analysis. However, the principal assumptions of sampling remain the same, as do the possibilities for using ISSA to support qualitative inquiry. Because qualitative studies are typically less concerned with macro-level analyses (although there are exceptions) we focus on individual persons as a unit of analysis for rhetorical clarity.

[5] Some qualitative methods, such as Bayesian analysis, do not emphasize determination of true effects. Like all positivists and instrumentalists, some quantitative methodologists at the turn of the century regarded theoretical constructs as inherently unobservable and therefore not accessible to scientific inquiry. However, our focus here is on integrating post-positivist quantitative and qualitative methods; hence we center our discussion of quantitative inquiry on methods for statistical generalization through sampling.

[6] The law of large numbers means that larger samples reduce the variability of estimates: large samples produce estimates that provide more consistent information about the population as a whole. For example, the ratio of heads to tails in 1000 flips of a fair coin is likely to be closer to 50:50 than the ratio of heads to tails in 10 flips. More generally, the sampling variation—or change in a statistic's value from sample to sample—is reduced as sample sizes increase. Therefore small differences in the characteristics of two different samples (or two different groups within a sample) are easier to distinguish from random variation in large samples.

[7] Traditions and methods of qualitative inquiry differ in the extent to which they emphasize this point in the operationalization of research, but even grounded theory, which has been accused of being "objectivist" by some of its critics (Charmaz, 2000), is defended by its founders as incorporating this fundamental premise (Glaser, 2002).

[8] The term is originally Ryle's (1971), from whom Geertz borrowed the phrase in his well-known essay of the same name.

[9] There are other statistical techniques, such as applied longitudinal data analysis (Singer & Willett, 2003), for addressing data that consist of repeated measures of subjects in an educational context. However, such approaches can still suffer from problems associated with collapsing data across subjects and reshaping the underlying research question described for more typical quantitative analyses in the following paragraphs.

[10] The researchers would include only 11 dummy variables. The variance of a 12th dummy variable with 12 subjects would be accounted for by the first 11.

[11] Strictly speaking, we would say that the model might show that "students were 4.6 times more likely than not to develop explicit mathematical understanding in a design episode when they engaged in a participant framework of design than when they did not engage in a participant framework of design," and that "students were 2.2 times more likely than not to develop explicit mathematical understanding when they encountered expressive obstacles than when they did not encounter expressive obstacles."

[12] In the example presented, for instance, there is a clear sequence in the learning. Any analysis would have to account for temporal information in the model so as to preserve the assumption of exchangeability. Simply assuming invariance over time would remove the ability to represent the pattern over time.

[13] Although the process of representing observations in exchangeable units of analysis is similar to controlling for confounding variables in a traditional quantitative analysis, exchangeability does not therefore imply de-contextualization. An assumption of exchangeability acknowledges that similar causal mechanisms may be reasonably expected to be at work across observations, thus making it possible to look at repeated observations of the same context in an analysis. We examine this issue in more detail in our Discussion.

[14] Exchangeable observations are necessarily spherically symmetric or permutationally invariant (Kingman, 1978). Independent observations are spherically symmetric with intercorrelations of zero. Thus, as Kingman argues, "a series of observations whose stochastic structure is unaltered by permutations… can always be regarded as a sequence of independent variables with a common distribution function" (p. 195), and all traditional inference tests which assume independence are also valid under the weaker assumption of exchangeability.

[15] It would be impossible to recreate the same experiment with the same group of students—just as one can never step into the same river twice—because on any subsequent iteration, the students would have been changed by previous experience of the experiment. However, results obtained using ISSA show, in effect, that had we gathered more data—or if we could travel through a science fiction time warp and repeat the experiment—we would predict that a similar pattern in the data would be observed. In referencing a population that can not be sampled in theory or in practice, ISSA is similar to Neymans, use of counterfactual situations to assess the effects of a treatment by generalizing to a hypothetical population—although, of course in ISSA the goal is to generalize within the original sample (Neyman, Dabrowska, & Speed, 1923/1990; Serlin, Wampold, & Levin, 2003).

[16] The original image is from http://www.jadecat-studios.com/t_stripes2.jpg.

[17] Additional images from http://www.bahti.com/textiles.html.

[18] ISSA's utility will vary depending on the particular technique of qualitative analysis being used. It is often better suited to case-focused analysis (Weiss, 1994) than to portraiture (Lawrence-Lightfoot & Davis, 1997), for example.

REFERENCES

Charmaz, K. (2000). Grounded theory: Objectivist & constructivist methods. In N. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research, 2nd edition* (pp. 509–535). Thousand Oaks, CA: Sage.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences, 6*(3), 271–315.

Creswell, J. W. (2002). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

Draper, D., Hodges, J. S., Mallows, C. L., & Pregibon, D. (1993). Exchangeability and data analysis. *Journal of the Royal Statistical Society, Series A (Statistics in Society), 156*(1), 9–37.

Geertz, C. (1973a). The growth of culture and the evolution of mind. In *The interpretation of cultures: Selected essays* (pp. 3–30). New York: Basic Books.

Geertz, C. (1973b). Thick description: Towards an interpretive theory of culture. In *The interpretation of cultures: Selected essays* (pp. 3–30). New York: Basic Books.

Glaser, B. G. (2002). Constructivist grounded theory? *Forum qualitative sozialforschung (Forum: Qualitative social research), 3*(3), Retrieved December 6, 2004 from http://www.qualitative-research.net/fqs-texte/3-02/03-02glaser-e.htm.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine De Gruyter.

Goodman, N. (1978). *Ways of worldmaking*. Indianapolis, IN: Hackett.

Kingman, J. F. C. (1978). Uses of exchangeability. *The Annals of Probability, 6*(2), 183–197.

Lawrence-Lightfoot, S., & Davis, J. H. (1997). *The art and science of portraiture*. San Francisco, CA: Jossey-Bass.

Lindley, D. V., & Novick, M. R. (1981). The role of exchangeability in inference. *Annals of Statistics, 9*, 45–58.

Mislevy, R. (1996). *Evidence and inference in educational assessment (CSE Technical Report 414)*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

National Research Council, Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Neyman, J., Dabrowska, D. M., & Speed, T. P. (1923/1990). On the application of probability theory to agricultural experiments: Essay on principles. Section 9. *Statistical Science, 5*, 465–472.

Ryle, G. (1971). The thinking of thoughts: What is 'le penseur' doing? In *Collected papers* (Vol. 2). New York: Barnes and Noble.

Serlin, R., Wampold, B., & Levin, J. (2003). Should providers of treatment be regarded as a random factor? If it ain't broke, don't "fix" it: A comment on Siemer and Joormann. *Psychological Methods, 8*(4), 524–34.

Shaffer, D. W. (1997). Learning mathematics through design: The anatomy of Escher's World. *Journal of Mathematical Behavior, 16*(2), 95–112.

Shaffer, D. W. (2002). Design, collaboration, and computation: The design studio as a model for computer-supported collaboration in mathematics. In T. Koschmann, R. Hall & N. Miyake (Eds.), *Computer support for collaborative learning 2* (pp. 197–222). Mahwah, NJ: Lawrence Erlbaum Associates.

Shaffer, D. W. (2003). When Dewey met Schon: Computer-supported learning through professional practices. In C. McNaught & D. Lassner (Eds.), *Proceedings of the World Conference on Educational Media, Hypermedia, and Telecommunications 2003 (Ed-Media)* (pp. 2197–2205). Honolulu, HI: Association for the Advancement of Computers in Education.

Shaffer, D. W. (2004). When computer-supported collaboration means computer-supported competition: Professional mediation as a model for collaborative learning. *Journal of Interactive Learning Research*, *15*(2).

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Strauss, A. L., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publications.

Weiss, R. S. (1994). *Learning from strangers: The art and method of qualitative interview studies*. New York: The Free Press.

Yu, C. H. (2003). *Misconceived relationships between logical positivism and quantitative research*. Retrieved June 8, 2004, from http://www.aom.pace.edu/rmd/2002forum.html

AUTHORS

DAVID WILLIAMSON SHAFFER is Assistant Professor, University of Wisconsin, Madison, Department of Educational Psychology, 1025 West Johnson Street, Madison, WI, 53706; dws@education.wisc.edu. His area of specialization is how computers and other new media change the way people think and learn.

RONALD C. SERLIN is Professor, University of Wisconsin, Madison, Department of Educational Psychology, 1025 West Johnson Street, Madison, WI, 53706; rcserlin@wisc.edu. His areas of specialization are nonparametric statistics, multivariate statistics, and the philosophy of science and statistics.

Manuscript submitted March 3, 2004
Revision received June 17, 2004
Accepted September 16, 2004