International Journal of Qualitative Methods

INTERNATIONAL INSTITUTE FOR QUALITATIVE METHODOLOGY (IIQM)
UNIVERSITY OF ALBERTA

Article

Insight, Inference, Evidence, and Verification: Creating a Legitimate Discipline

Keynote Address for the II Congreso Iberoamericano de Investigación Cualitativa en Salud Madrid 22-25 de Junio de 2005

Janice M. Morse
University of Alberta
Edmonton, Alberta, Canada

Abstract

Challenges to the subjective nature of qualitative inquiry, recent interest in evidence-based practice, and the advances in mixed-method design have all contributed to the value and utilization of qualitative inquiry. The author is concerned, however, that the value placed on evidence and the agenda of qualitative-quantitative mixed-method design will devalue the role and contributions of insight and inference in our research. In this address, she argues that research using insight and inference can and must be used and valued in qualitative inquiry. Insights invariably arise from single instances, exemplars, or single-case studies, albeit often, but not always, verified in other instances. But the knowledge gained from insight might make a greater contribution to scholarly study than replication and verification, for without insight, research can be mundane and obvious.

Keywords: standards, abstraction, verification, evidence, interpretation

Introduction

Qualitative research has developed in the shadow of quantitative inquiry. While we have often associated the popularity of quantitative research with the development of computers, which provide these researchers with the ability to analyze vast data sets and with confidence in the accuracy of their results, I am not sure that is the whole story. We have grumbled at the stronghold of quantitative researchers in academia, in granting agencies and journal review boards, and blamed quantitative researchers for the difficulties we have had in establishing the acceptability of qualitative inquiry. But today, I am suggesting that we are at least partly responsible for our difficulties, for the lack of self-esteem held by qualitative researchers, and for the lack of confidence we ourselves hold in the generalizability and utilization of our results. This problem stems from our very own hands—for it is we who establish criteria and standards for qualitative research, for reliability and validity. And it is we who made these rules and chose to model them after quantitative criteria. We must, therefore, accept some responsibility.

How did we get in this mess?

To answer one question with another question: How many of us were initially "trained" in quantitative inquiry? Most of us, probably. In North America, even today, quantitative inquiry is considered essential in undergraduate and graduate programs; qualitative inquiry is considered a luxury, something to be added to the curricula only if there is space and if there is time.

For us—for established qualitative researchers—this means that we first had quantitative principles burned into our research brains; then this knowledge was overlaid with qualitative principles, rather as an afterthought. For us it means that we must become comfortable and confident with our qualitative research, in light of this ingrained quantitative baggage. Thus, is not easy to become confident with qualitative principles. Qualitative principles do not follow quantitative principles logically; on the contrary, they are often discrepant, and the rationale for these discrepancies has not been made explicit. At the same time, to add to our discomfort, we exist in a world that does not value, and is constantly questioning, the nature, role, and function of qualitative knowledge.

Therefore, it should be no surprise to us when we look at developing standards for qualitative research we find that we are emulating the dimensions and components of quantitative standards—albeit often defensively. Even our agenda is misguided, in that we are seeking sets of single criteria to apply to an extremely diverse range of methods—methods based on different assumptions, using different kinds of knowledge, at different levels of abstraction, and accounting for different phenomena. Yet, we are setting this agenda; we are doing this work. And we wonder why there is discord and disagreement among us.

So, today I am suggesting that we must start anew, reexamining what we are doing and why, starting with the assumption that together we are creating a new discipline, one that will make a new, different, and significant contributions to developing knowledge. I will make my case in the context of four areas of qualitative research: insight, inference, evidence, and verification.

Insight

Insight is the neglected and overlooked mechanism, the Cinderella of qualitative inquiry, rarely consciously used, seldom boasted about, and almost never valued as one of the major processes in qualitative analysis. We never read, " . . . and then I had an idea." On the contrary, our research reads as though we developed our stunning results systematically and logically, directly from our data.

It is a fact that insight invariably arises from a single instance, exemplar, story, or case study, then often (but not always) verified in other instances. These instances may be exactly the same or different examples of the same concept, which can be verified in the past within data already collected, or recognized as future occurrences. But the knowledge gained from insight, or the contribution of insight to our study, is more significant than ascertaining rigor through saturation, replication, and verification, for without insight, our research can be mundane, obvious, and atheoretical.

What do we know about insight? We know that it must "fall on fertile ground" or it may be missed. That is, for an insight to

occur, the researchers' minds must be ready. They must have much knowledge about what they are observing or hearing, know and be able to link it to relevant literature, and be able to think conceptually and to link seemingly unconnected events, representations, and ideas.

How do we foster insight? How do we make qualitative inquiry as theoretically rich as we can without going beyond the borders of reasonable theorizing with our project? The answer to the first, making our research theoretically rich, is simple—give ourselves and others permission to use insight and the confidence to do it well. That said, we must also prepare the "fertile ground"—we must use the library, read everything possible before going into the field, so that we will be able to recognize a "wheel," for instance, if we see one. We must become immersed in our data, have good-quality data, well-scoped data, and seek saturation (Morse & Richards, 2002).

But there is something else: We know that the practice of "researcher-as-the-instrument" is fundamental to qualitative inquiry, but we seldom discuss why. Apart from the literature on establishing trust with our participants, we appear embarrassed by this use of self that is inherent in our research process. We defensively counter criticisms of bias and other concerns with detailed strategies for maintaining neutrality to prevent threats to validity, and so forth. In these debates, we almost never discuss the fact that insight is crucial to understanding "what is going on," for seeing the implicit, for uncovering, for interpretation, and for developing strong concepts and theories; in other words, for building excellent qualitative inquiry. Our fieldwork/data collection and analysis is far from the technical and rote task that quantitative researchers enjoy, yet ironically, we still emulate quantitative research and its rules for the conduct of research in the clinical setting.

(Note, I am selecting to discuss the clinical setting in this example, because it exemplifies the problem we may face using a quantitative research model.)

One norm that quantitative researchers have is using the clinical areas in which they work on a daily basis, and their own patients, in their research. This research may be a clinical drug trial, with patients recruited and randomly assigned to a treatment or to a control group. Scientific concerns for coercion and bias, and ethical concerns for harm are eased by having a third party recruit patients into the trial and procedures for breaking the treatment/medication code should untoward effects of the experimental drug occur.

With the exception of the concern for coercion in recruitment, the threats to quantitative research (or the subjects) are not pertinent to the types of research designs used by qualitative researchers. But this does not mean that qualitative research can be conducted in the researcher's own clinical practice without concerns' being addressed!

The greatest threat in using one's own clinical areas and patient population in qualitative inquiry is that it results in the loss of insight. Anthropologists initially conducted research in cultures in which they were strangers—simply to maximize insight, recognizing how the sharing of one's own culture with participants inhibited the identifying and understanding of values and beliefs. When one works in one's own work area, insight is lost. Habits and norms dull one's sense of seeing whatever is really happening, and stifle the observations that may slap a stranger in the face. One has limited sensitivity, accepting the extraordinary as ordinary.

I recall reporting to trauma room nurses on comforting strategies I saw them using with distraught patients. When I finished, they looked at me in astonishment. "We do do all of those things," they said. "But we do not consider them a part of our work!" These were nursing actions, such as talking the distressed patient through a painful procedure—nursing actions that I later came to consider crucial for

the survival of the patient, as they increased patient endurance and prevented the patient from fighting caregivers!

Therefore, if you are interested in a problem that has come from your practice, do not conduct your research in your own clinical setting, such as your own ICU. Change hospitals, and use an ICU in which you are a stranger.

Interviewing one's colleagues and friends also alters the type of information obtained. These interviews will not contain everyday information with which you are both familiar; thus, significant issues may be omitted from your data set.

But there is another factor that must be discussed with ethics committees, which, to my knowledge, is not considered during review. If you are conducting qualitative research on your own patients, the research information obtained crosses over as clinical data, and similarly, clinical information may also be considered research data. But the patient may be consciously aware of when he or she is a patient and when he or she is a research participant. Removing this distinction between being a patient or a research participant has implications for violating informed consent, thus violating the patient's rights and placing the researcher/nurse in an impossible conundrum. The patient may not be aware of whether you are treating the information as a researcher or as a clinician, and the researcher/clinician may not know which "data set" the information should be recorded in. Both are confidential data sets, but different teams of employees have access to each data set. Thus, what does a researcher do, when the clinical information is needed for the research component, and vice versa!

But it gets worse. If the researcher is also a counselor, when does the counseling cease, or when is it used, in a research interview?

From this single example, it is clear we must develop our own, independent set of guidelines.

Inference

Inference allows us to make linkages; inference allows us to recognize new instances of a case; inference allows us to categorize—to put somewhat similar things together—in other words, to create categories and themes. But most important, inference allows us to link concepts, to create theory, to apply our research results, and to move forward. We do this without requiring measurement and p values but by using qualitative indicators and identifying characteristics, by verifying each step in the process, and by using common sense.

In quantitative inquiry, inference is used in the initial stages of constructing a conceptual framework, and inferences are stated as assumptions or in the research process as hypotheses to test these inferences. In qualitative inquiry, inference is used on an ongoing basis during data analysis. Closely allied to insight—in fact, one could define it as a planned insight—inferences are developed and "checked out" on an ongoing basis during analysis until one is convinced and the inference is no longer a tentative conjecture but becomes an "analytic fact." It is this process of verifying inferences that enables abstraction of the emerging theory, and the development of cohesive and congruent results.

Evidence

Qualitative researchers have been shifting uncomfortably on the perimeters of the evidence-based movement. Quantitative researchers have convinced us (among others) that quantitative measures are the only appropriate "currency" for evaluating evidence; to demonstrate what is better, larger, more effective; and so forth.

The new standards for evidence have developed rapidly in the past decade, in the form of the standards initially set by Cochrane to determine efficacy (Morse, 2005). Fuelled by the pharmaceutical industry, these standards, most helpful for determining the efficacy in clinical drug trials, have little relevance for the type of questions that interest qualitative researchers—yet these standards have been used by medical granting bodies as criteria to award funding, whether relevant or not. This is a political problem confronting qualitative inquiry, which we must somehow resolve.

Nevertheless, a group of researchers under the leadership of Jenny Popay from the United Kingdom, have been working for a number of years to move qualitative inquiry into Cochrane reviews. Last year, I wrote that this work has made little headway, but this year I take back my words somewhat. I think this group has made important contributions to the visibility of qualitative inquiry. However, the fact remains that qualitative inquiry does not measure things, just as quantitative research is clumsy at determining meanings. We need to develop alternative criteria for evidence, using other standards and techniques, such as logic, qualitative evaluation, and common sense.

Quantitative researchers' preoccupation with measurement is evident in health care in the area of pain. Despite many years of struggling to quantify pain using physiological indices, researchers remain stymied. We do not have a dipstick to measure experienced pain. But clinicians have achieved the stunning feat of teaching patients to quantify their own pain experience. Over the past decade, clinicians have been teaching patients to objectify their pain experience by reporting pain, not in complex facial grimaces and colorful adjectives but by ranking their pain using a numerical scale, with 10 being the worst pain ever experienced (VNS, the 11-point Verbal Numeric Scale) (Herr, Spratt, Mobily, & Richardson, 2004). Elsewhere I have argued that removing the patient's prerogative of graphically describing the pain removes the sharing of the pain experience—the transmissions of the compathetic response (Morse & Mitcham; Morse, Mitcham, & van der Steen, 1998)—with the caregiver, as well as saving a remarkable amount of time. But is the amount of pain experienced by the patient communicated and appreciated by the clinician when a number is used? In an editorial, I asked, imagine if such techniques were used in everyday life:

> "How much do I love you?"
> "Oh, 4.5." (Morse, 2005b, p. 144)

Now your laughter is my evidence that qualitative data are richer! Let us add that to the list of things to do: Develop criteria of evidence using logic, qualitative evaluation, common sense, and laughter!

Verification

One of the most debated areas for developing standards is in the area of "reliability and validity." Some researchers recommend participant verification of transcripts on other forms of member checking; others insist on determining inter-rater reliability during coding, while others—and put me in this camp—complained that such strategies stifled creativity and insight in research.

In quantitative research we are taught that reliability and validity are serious concerns, and without adequate levels of each of these, our research is worthless. But in qualitative inquiry, we are concerned not with measurement but with description and meaning; hence, reliability and validity take on a different role.

Three features of qualitative inquiry are important: First, if the purpose of the study is to describe, then some form of reliability checking is important. This is true with research that is descriptive, with semistructured questionnaires, and with various forms of data transformation in QUAL-QUAN mixed-

method designs (Morse, Niehaus, Wolfe, & Wilkins, in press). It is necessary when accuracy is paramount, as in conversational analysis, and for microanalytic observational work, such as facial-coding or behavioral description. I refer to this as direct data, for it must represent the actual phenomena very closely.

Second, strategies for ensuring reliability may also be used with data in which more error is tolerated. This may be research that examines narratives and perceptions and emotional responses, but the participant's perception of the event is generally considered in qualitative inquiry to be accurate, and thus, their perceived experience as it occurred is considered valid. These semidirect data include participants' reports of what happened, recalled events and conversations, and so forth. They are approximations, and some variation occurs between participants, because the reports are perceptions of occurrences. In anthropology, these discrepancies are considered the Rashomon effect (Heider, 1983).

The last type, indirect data, are data that are inferential and abstract (not concrete), consisting of signs (such as nonverbal behavioral indicators of meaning, sarcasm, covering), signals (such as the use of shadowed data or the participant's reporting of others perceptions), and symbols (representations, such as metaphors)—types of data that, by definition, may or may not be valid, may be right or wrong, but these questions of accuracy are of little importance—as blasphemous as that sounds. Yet, these data reveal the implicit, are symbolic, contribute to concept and theory development, and make good qualitative data stimulating, surprising, exciting, and innovative. As people, our worlds extend beyond the concrete, and this must also be included in our research.

Does this mean that qualitative researchers cannot go wrong? For goodness sake no! It means that with different types of data—and with different methods—there are different means of determining rigor. If, for instance, we used inter-rater reliability techniques with the third type of data—I call them indirect data—we would be wasting our time—for there is no "right" and consistent answer to be obtained. Or if we used it with the second type—with semidirect data—we would be squelching the creativity, keeping our results close to the data and preventing abstraction. To use an analogy from an American children's story (O'Brien, 1971), our results would be "Perfectly healthy but dead"!

Thus, with indirect data, other criteria for determining rigor must be used. For instance, if we are using a method that creates theory, then we use criteria for evaluating theory (Glaser, 1978; Morse, 1997) or criteria from philosophy using rules of logic.

I want to mention one last thing about verification: Because qualitative inquiry is verified in the process of data collection and analysis, good qualitative inquiry must be verified reflexively in each step of the analysis. This means that it is self-correcting—inadequate or poorly supported constructions are not supported and "fall out" of the analysis. In this way, qualitative inquiry, properly conducted, is self-correcting and rigorous, and the results are strong.

Creating a new discipline

Because of the differences between qualitative inquiry and quantitative inquiry, we must not and cannot borrow rules from quantitative inquiry for the conduct of inquiry, assessing ethical risks, or ascertaining rigor. The agenda of qualitative inquiry invites us to consider new dimensions for the consideration of ethics, and new threats to validity, and unimagined processes to determining reliability. But until qualitative inquiry is mainstream, and until all of these problems have been clearly identified and explained, we cannot use quantitative standards to evaluate qualitative research. Guba and Lincoln's (Guba, 1981; Guba & Lincoln, 1981; Lincoln & Guba, 1985) introduction of the concept of trustworthiness is, I think, only a partial solution to our problem. I believe their rejection of reliability and

validity and replacing it with trustworthiness, a carte blanche term for use with all kinds of qualitative inquiry, is simplistic.

The framework introduced here, considering direct, semidirect, and indirect data is more useful than trustworthiness (although it could feasibly incorporate it and tolerates various modes of ascertaining of rigor). We must focus on the type and role of data, and develop appropriate strategies for ensuring quality.

Why are we bothering to develop these new methods? It is not an exercise that we do just to disturb the status quo. It is something that we are doing because of the product of qualitative inquiry.

- With the use of qualitative insight, we are creating a new perspective.
- With qualitative inference, we are contributing new theories.
- With our development of qualitative evidence, we will contribute new knowledge.
- By the appropriate use of verification, we will develop new practices.

These changes will not come quickly, and they will not be easy. But a change in the way we do social science research must occur.

While we may consider the qualitative and quantitative divide, not categorically but as a continuum, there is no doubt that qualitative and quantitative paradigms do do things differently. These differences are so apparent that my wish is for qualitative research to be considered a different discipline, and not reviewed by committees with no qualitative expertise nor appreciation. We need our own granting committees with a percentage of funding allocated for qualitative inquiry, rather than expecting quantitative researchers' to share their funding—for there is ample evidence that quantitative committees are uncomfortable with and lack of understanding of qualitative strategies, principles, and methods; without such fair treatment, our progress with be limited.

But this methodological work for developing the discipline is on our shoulders and is not the purview of quantitative researchers. We must first see ourselves as a distinct discipline, making a unique and important contribution to knowledge.

Thank you

References

Glaser, B. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology Journal 2*, *9*(2), 75-91.

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.

Heider, K. (1983, Spring). The Rashomon effect. *Association for Social Anthropology in Oceania Newsletter*, pp. 10-11.

Herr, K. A., Spratt, K., Mobily, P. R., & Richardson, G. (2004). Use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. *Clinical Journal of Pain*, *20*(4), 207-218.

Lincoln, Y., & Guba, E. (1985). Establishing trustworthiness. In Y. Lincoln & E. Guba (Eds.), *Naturalistic inquiry* (pp. 289-331). Newbury Park, CA: Sage.

Morse, J. M. (1997). Considering theory derived from qualitative research. In J. M. Morse (Ed.), *Completing a qualitative project: Details and dialogue* (pp. 163-188). Newbury Park, CA: Sage.

Morse, J. M. (2005a). Beyond the clinical trial: Expanding criteria for evidence [Editorial]. *Qualitative Health Research, 15*, 3-4.

Morse, J. M. (2005b). Decontextualized care [Editorial]. *Qualitative Health Research, 15*, 143-144.

Morse, J. M., & Mitcham, C. (1997). Compathy: The contagion of physical distress. *Journal of Advanced Nursing Practice, 26*, 649-657.

Morse, J. M., Mitcham, C., & Van der Steen, W. J. (1998). Compathy or physical empathy: Implications for the caregiver relationship. *Journal of Medical Humanities, 19*(1), 51-65.

Morse, J., Niehaus, L., Wolfe, R., & Wilkins, S. (in press). The role of theoretical drive in maintaining validity in mixed-method research. *Qualitative Research in Psychology*.

Morse, J. M., & Richards, L. (2002). *Readme first for a user's guide to qualitative methods*. Thousand Oaks, CA: Sage.

O'Brien, R. C. (1971). *Mrs. Frisby and the Rats of NIMH*. New York: Athaneum.