Interpretation and inference: towards an understanding of methods

Geoff Norman

Received: 6 October 2010 / Accepted: 6 October 2010 / Published online: 21 October 2010
© Springer Science+Business Media B.V. 2010

Last week I was chatting to a colleague about an introductory methods course we were working on. We were talking about the unit on “The Research Question, and how this differs between qualitative and quantitative research. He then made an offhand remark along the lines of, But then, the qualitative—quantitative debate is over anyway. Everyone is doing mixed methods research.” I was greatly relieved to find that this debate, which has rumbled on since I was a student, is finally resolved. But since I wasn’t really sure how mixed methods had won the final round, I asked him, somewhat naively, just what mixed methods were. He explained that this was when you did a survey after your qualitative study to get some check marks on Likert scales so you could do statistics on it and get a p-value. When I tried to explain to him that conducting a survey was only one of the most basic methods of quantitative research, and indeed, many quantitative researchers do not accept the findings of most surveys, he simply did not understand. This reminded me of the last iteration of “mixed methods” I heard where you designed your survey questionnaire, then conducted a “focus group” in which you handed out your questionnaire and asked people what they thought of it. It will surprise no one that the first version was advocated by a person who is primarily a qualitative researcher, and the second by someone who bears allegiance to quantitative research. The first trivializes quantitative research; the second trivializes qualitative.

My colleague went on to propose that all educational researchers should be well versed in both methods so they can move easily from one to the other. That seems to me equivalent to demanding that all undergraduates should be competent in quantum mechanics and theology. It would be nice if each should be sufficiently familiar with the assumptions, epistemologies and methods of the other to be an appreciative observer, but each discipline is sufficiently complex and multi-faceted to claim full time attention of anyone working toward achieving competence. Indeed, even within quantitative methods, there are many highly specialized subgroups—psychometrics, cognition, personality, and so on. The same, I suspect, holds true, in qualitative; I presume battles rage between ethnomethodologists, grounded theorists, and Foucauldians, just as they do between

G. Norman (✉)
McMaster University, 1200 Main St. W, Hamilton, ON, Canada
e-mail: norman@mcmaster.ca

experimentalists and psychometricians (Cronbach 1957). But then, I don't understand qualitative methods well enough to be sure on this point.

Is the continuing debate between qualitative and quantitative simply a question of increasing the understanding of each to the other? I fear not. I am continually struck by the fundamentally different epistemologies. Recently this resurfaced as an issue of the degree to which each side allows for or, in fact, rewards interpretation. Interpretation—informed interpretation—is at the heart of all science. The fundamental notion that the facts are value-free and “speak for themselves” is a naive perspective dismissed by all philosophers of science. As Mary Budd Rowe said, and I have repeated many times:

Science is a special kind of story-telling with no right or wrong answers, just better and better stories.

To considerable degree, the scientist is in the business of writing good stories, based on good data. The facts are only the starting point; how they are interpreted involves theories and metaphors. But it seems to me that in the translation from data to story, qualitative and quantitative begin with different “meta metaphors” about how interpretation proceeds, and that ultimately lead to different strictures and canons. The quantitative folks like to think of the journey from facts to knowledge like the parades of Red Army regiments on May Day, where the facts move in complete and precise relation to one another in a straight line from question to conclusion—all regimented, all precise, all linear. By contrast, the qualitative metaphor seems more like Swan Lake performed by the Kirov Ballet; harmonic and lyrical movement through crescendos and glissandos, all beautifully orchestrated and choreographed.

Regrettably reality is seldom like this. Too many of our quantitative papers read like a platoon of recruits at boot camp. Everything is moving in generally the same direction, but the movement from beginning to end is sloppy and full of stumbles and detours. Any sense of precision and direction is lost. As to qualitative, to me it often resembles postmodern ballet, where the orchestration is atonal discordant, the dancers appear in their rehearsal tights, and each interprets away with little acknowledgement of the others. The sense of elation and joy that comes from the soaring orchestration and elegant coordinated movements of the dancers in the Nutcracker or Eugene Onegin is conspicuously absent. The platoon is comical; the modern ballet is incomprehensible.

Contrast the following quotes, both from the discussion section of articles in recent issues of *Medical Education*. The first was an ethnographic study; the second, a psychometric paper:

Rather, on a more general level, I wish to suggest that teaching strategies that encourage independent goal setting and research on the part of the learners can mirror the rearrangement of power relations and responsibilities that patient-centredness also advocates. I draw attention to the fact that, within the undergraduate curriculum, the way in which tutors problematise their role in supporting students' autonomy in learning has clear relevance to the ways in which students construct their understanding of, for example, shared decision making in the clinical consultation. Integrated with the other pedagogical approaches accommodated by the undergraduate curriculum, more collaborative pedagogical strategies may help prompt students to reflect critically on authority, power and responsibilities in medical practice. (Donetto 2010)

The validation of the D2 predictions on a totally independent dataset (Table 6) shows that, despite the different statistical approaches adopted by Formula 11 (for predicting

future G) and Formula 7 (for calculating actual G), there is very close correspondence in the G estimates, indicating that these formulae are accurately identifying the variances caused by an unbalanced, uncrossed and fully nested design. Furthermore, the D calculations in the case studies indicate that there may be a tendency to over-sample patients, perhaps in the belief that patients' responses are less reliable (more varied) than colleagues. Table 3 shows that, on average, patient responses tend to be less varied than colleague responses (0.14 versus 0.18 for GMC study data, 0.17 versus 0.35 for GP study data) because of the usual and limited role that patients have in their interaction with doctors. (Narayanan et al. 2010)

The point of the metaphor is to highlight that interpretation has very different meanings for the two camps. To the quantitative researcher, interpretation is almost a bad word. We like to think that the data speak for themselves, and we venture timidly, indeed often far too timidly, into the area of interpretation. Students are frequently cautioned to "don't go beyond your data" and all conclusions are well guarded by safe modifiers like "suggests", "perhaps" and "may" with the inevitable call for more research before any conclusions are drawn. When interpretation exists at all, it is generally the last step in the linear and predictable procession from Introduction through Methods to Discussion, right after the section on Study Limitations. There are two problems with this approach. First, in contrast to much qualitative writing, far too many quantitative papers are so formulaic and pedantic that they are about as exciting to read as the Yellow Pages. But more important, far too often the discussion section is restricted to reiterating what was found, and there is too little effort to link the study findings to theories or existing literature and to show how the paper contributes to elaborating the theory.

By contrast, I continue to be amazed and awe-struck by the writings of qualitative researchers. Sometimes even the critiques of qualitative papers that I read as editor seem worthy of publication as examples of fine literature. None of us in medical education will ever win a Nobel Prize, but some qualitative writings look like they deserve the Booker Prize. But therein lies the other edge of the double-edged sword. It seems to me that the thread between data and interpretation is frequently thin indeed. Although qualitative methodology is all about placing defensible interpretation on observations using strategies like triangulation and saturation to ensure that the interpretations are valid and verifiable, nevertheless, when a paper offers a selected quote from a transcript as evidence for an observation, it is impossible for the reader to independently verify the interpretation. I was recently discussing a qualitative paper with a colleague, where the draft was very light on both methodology and original data. He said, "It's a superb essay." I responded, "Yes it is. But it's supposed to be a study". It was really the case that the paper could have stood on its own with all the cited comments omitted, with very little loss of veracity. Or conversely, the presence of the data added very little to the paper. It is hard to imagine a quantitative paper being published that omitted the methods and results section, and went directly from background to discussion.

It seems to me that each camp can learn from the other. Too often quantitative research is ignored because, in the attempt to maintain caution, the contribution of the particular work to the larger literature is absent. Too often qualitative research is ignored because the linkage from original data to conclusions is left unclear, so the findings are viewed as more like opinion than evidence. A middle ground is possible and it may ultimately lead to greater understanding on both sides.

References

- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 61–84.
- Donetto, S. (2010). Medical students' views of power in doctor-student relationships: The value of teacher-learner relationships. *Medical Education*, *44*, 187–196.
- Mennin, S. (2010). Self-organization, integration and curriculum in the complex word of medical education. *Medical Education*, *44*, 20–30.
- Narayanan, A., Greco, M., & Campbell, J. (2010). Generalizability in unbalanced, uncrossed and fully nested designs. *Medical Education*, *44*, 367–378.