

## Sample sizes, scoops and educational science

Geoff Norman

Received: 29 October 2010 / Accepted: 29 October 2010 / Published online: 25 November 2010  
© Springer Science+Business Media B.V. 2010

It seems that lately I am being consumed by an obsession with philosophy of science, and the nature (epistemology) of research in medical education. Perhaps it is a natural consequence of a sabbatical (just finished), which often leaves you with too much time on your hands to think deep thoughts. So this is the third editorial in the series to examine an aspect of our research game. Hopefully, not too much repetition.

Three seemingly unrelated personal anecdotes, both recent and ancient:

(1) Recently I received an e-mail from a medical student along the lines of:

“Dear Dr. Norman. I have a study I want to do. I’ve met with the chair of the undergraduate program to discuss it, and he suggested I meet with you to do a sample size calculation”.

I wrote him back indicating that, while I can meet with him, meetings of this sort may start out as sample size calculations but rarely end up that way, as I may have a number of suggestions about how to do the study differently. What I wanted to say, but didn’t, was A) I’m not your neighbourhood sample size monkey. Go and Google something to get a sample size, why don’t you? B) Why do you think that your question, as yet unmentioned, deserves the resources needed to conduct a study? Have you looked at the literature? Have you read papers on it? Why do you think your bright idea deserves any attention until you do this?

On reflection I realized that I was being overly sensitive, and this was not the way I would have gone about it a few years ago. I would have welcomed him into my office; thrilled that he showed the initiative to get interested in education research. We would have talked about sample selection, research design, measures, analysis approaches and so on. We would NOT have talked about whether the question was worth asking. My role was to encourage research in education—ALL research in education. I may still do that, now that I’ve calmed down.

(2) Last week, I was reviewing study questions generated by students in our introductory epidemiology course. One student wanted to solve all the problems of childbirth in the developing world; another wanted to do a horrendously complex study of remissions in

---

G. Norman (✉)

MDCL 3519, McMaster University, 1200 Main St. W., Hamilton, ON L8N3Z5, Canada  
e-mail: norman@mcmaster.ca

multiple sclerosis. Neither had made any attempt to put their study in context, to see what was already known about the area, to show how their question was a logical next step in understanding. It was not simply that the questions were unanswerable; it was also that there was no reason to presume that they were worth answering, at least from a scientific perspective. They might have been hugely important in terms of lives saved, anguish prevented, and such, but it was not clear if they represented any contribution to scientific understanding.

As I thought about this, I realized that the students were not at fault. The first week of the course was all about the Research Question, and how to design a question that could be answered. Not a question that *should* be answered; one that *could* be answered. Nowhere in the course, or in all the methodology books and review guidelines have I ever seen, is any discussion of how to ensure that the question was worth asking in light of existing knowledge.<sup>1</sup> It is just assumed that as intelligent observers of their community of practice (if they had one) or reading of the newspaper, they could easily produce a question worth answering. We just assume that a question is OK if we can figure out a credible way to answer it, and we put all of our efforts into good designs, good measures, good analyses—good answers, not good questions.

(3) When I was a graduate student and finishing off my PhD in physics, I realized that, whatever other skills I had acquired, one I never did figure out was how to take a book out of the library. I had got a PhD without ever checking out a book. While some will view this as one of many reasons why I never made it as a physicist, at the time it made perfect sense. As an undergraduate, my time was spent learning 400 year old Newtonian mechanics; there was no divergence of opinion in the textbooks about what equation was what. As a graduate student, it was either quantum mechanics, 50 years old, or, as I got going on my research, current research in the journals. In fact, while I did not take books from the library, a Friday afternoon ritual was to go to the library to check out the latest issue of the journals. As I flipped the pages, my heart raced and my palms sweated, for fear that I would be “scooped”.

“Scooped”? That’s something journalists worry about; someone got their story into print before they did. But as Richard Tiberius describes, this is also a common concern in natural science, with exactly the same meaning. If someone publishes the “story” of the doorway states in the neutron cross-sections of light nuclei (for this they gave me PhD), then I can’t. Replication has almost no place in physics. Systematic reviews are unheard of. If it’s done, it’s done and, with rare exceptions, it does not get redone. No one tries to argue that their neutrons are different from the other guy’s, or they measured the energy of their gamma rays with a different detector, or they came from a different conceptual framework. And no one would dare say, “The energy of mu mesons has not been measured in South Columbia before, so we did a study.”

What is the common element in these three stories? It all comes down to the role of the research question. As a science matures, the corpus of knowledge becomes more coherent and is shared by the network of scientists. Research questions arise from the shared knowledge of the community, not from whole cloth. The literature review is not just a bit of scholarly fluff to ensure that there is a half page of references at the end of the paper; it is the essential argument leading to why the question is worth asking.

There was a time when many papers in our field began with something like, “At our medical school we had a problem with...” Well, the times they are a’changin’ and this kind of rationale is no longer acceptable. At least two journals in medical education have

<sup>1</sup> With at least one exception. See Norman and Eva (2010).

now changed their submission guidelines so that authors must explicitly show how their study contributes to knowledge. Reviewers are asked to make direct comments on the unique contribution of the study. And on more than one occasion, I have turned down a paper that was perfectly adequate from a methodological perspective, but did not add new knowledge. These changes in direction, subtle though they may be, represent a maturation of the field, and may well signal that we must adopt a different way of doing business. However, subtle the changes, they are profound. Two clear implications:

(1) When I entered the field, we viewed that the optimal research team was a partnership between a practitioner or clinical educator, who knew, from her personal experience, the good questions to be asked, and an “educationist” like me, who knew the methodologies available to answer them. They had the questions; we had the methodological answers, clothed in arcane terms like analysis of variance, construct validity, regression. We joked about “the clinicians and the magicians”.

While the model is still prevalent and functional, as we mature, the role distinction changes, so that far more frequently, questions do not arise from individual experiences and observations, but instead emerge as the next step in a research program. And the program is typically led by an “educationalist” or more accurately, a behavioural or social scientist (Norman 2010). As one indicator of this change, the leaders of the field in the early days were often as not, clinicians—George Miller, Hilliard Jason, Howard Barrows, Victor Neufeld., Paula Stilman, Ron Harden (who still is). However, more recent “snapshots” of the field (Schmidt and Rotgans 2010) show that, for example, eight of the ten most cited authors in our journals are PhDs. This does not mean that there is no role for clinicians; indeed quite the opposite. Effective research will come from a much closer collaboration, so that each has a more intimate understanding of the other’s perspective, at a more fine-grained level.

(2) Recently some authors have argued (Regehr 2010; Mennin 2010) for a changing paradigm of educational research, where there is a minimal role for generalizable theories. For example, Regehr writes:

It [the science of education] is about exposing our underlying metaphors and assumptions, and examining the relative value of these metaphors and assumptions for interpreting the education issues that we are individually and collectively trying to address in our own local contexts. (p.37)

This discourse places value on local solutions as metaphors for other local solutions, and specifically discounts any attempt at generalization. To quote, “the science of education is not about creating and sharing better generalisable solutions to common problems.” Taken at face value, the “science of education” amounts to a random collection of individual demonstrations, not a coherent body of knowledge. Our role as educational scholars amounts to accumulating large collections of exemplars, then sifting through the demonstrations to find some solution in some setting that approximates our own. We become more like legal scholars, looking through the archives for precedents to apply to the current situation, than scientists. However, to the extent that the trends I identified above hold true, the opposite is the case.

Donald Campbell once said, “A study is not worth doing well if it’s not worth doing”. It has taken us a long time to understand and act on the profound meaning behind this simple sentence. But that time has now arrived.

## References

- Mennin, S. (2010). Self-organization, integration and curriculum in the complex world of medical education. *Medical Education*, *44*, 20–30.
- Norman, G. R. (2010). Forty years of medical education research: The actors and the play. *Medical Education* (in press).
- Norman, G. & Eva, K. W. (2010). Quantitative research methods in medical education. In T. Swanwick (Ed.), *Understanding Medical education: Evidence, theory and practice*. Chichester: Wiley-Blackwell.
- Regehr, G. (2010). It's NOT rocket science: Rethinking our metaphors for research in health professions education. *Medical Education*, *44*, 31–39.
- Schmidt, H. G., & Rotgans, J. (2010). The themes of medical education research 1988-2010: Content analysis of abstracts from five journals. *Medical Education* (submitted).