Mathematics Education Research Journal 1993, Vol. 5, No. 1, 19–22

Babies, Bathwater and Straw Persons: A Response to Menon

Sid Bourke
University of Newcastle

I have found writing this response to be a difficult task, as evidenced by my inability to resist the combination of cliches in the title. As I read Menon's article I found myself agreeing with much of what he had written, although sometimes I wondered why it was considered to be worth stating. Then Menon would take a more extreme line which had not really been justified by what had preceded it, and I found myself frustrated by the lack of continuity as much as by the extreme view itself. I will give some examples of what I found to be problems with Menon's position, based around the themes of (a) was it worth saying anyway; (b) the function of over-statement; (c) methodology and the role of theory in educational research; and (d) the proposed ideal world of educational research. In this response I have taken research in mathematics education to be entirely subsumed in educational research generally.

Was it Worth Saying?

Whereas stating the obvious may sometimes be a useful reminder, it is then not normally a good idea to go on to make a major issue of it. For example, Menon (1993, p. 4) quotes Tyler cautioning against the uncritical use of statistical significance testing (SST). Uncritical use of any research or analytic technique (or of anything at all for that matter) does not constitute good practice. Here I am stating the obvious, as a reminder. Menon is correct (and so was Tyler) but what is the point of going over old ground now? The assertion immediately following that "tests of statistical significance continue to dominate the interpretation of quantitative data in educational research" (Shaver, cited by Menon, 1993, p. 5) cannot be used to establish that all or any of these uses were uncritical.

The Function of Overstatement

Menon seems to be asking the reader to believe that there are only two, mutually-exclusive choices with respect to decision making: Fisher's significance testing or the use of one's "own informed judgement" (Gigerenzer and Murray, cited by Menon, 1993, p. 5). The informed judgement of eminent figures in a particular field should certainly be taken into account by researchers, but is that all the information they should gather? How eminent does one have to be to have one's judgement counted as evidence?

Again overstatement spoils a useful point being made on the frequent use of SST. Because of claimed misuse of SST ("uncritical and mechanical application"), Menon suggests that SST should not be taught to intending researchers. The views of Carver and of Coates from the 1970s (cited in Menon, 1993, pp. 12–13) are

presented as the main evidence of misuse. Even if these views of SST teaching were correct then, one could reasonably doubt that the same still holds. Has Menon looked at contemporary teaching of budding researchers? I teach quantitative research methods to graduate students. The basics of SST have a place in these subjects as do concerns such as the need for theory, the differences between statistical and educational significance, descriptive statistics, and estimates of effect size. I do not believe that my courses are very unusual in education faculties of the 1990s. One might have thought that better teaching of the place of SST in quantitative educational research would be a more reasonable suggestion for Menon to have made. But it seems that Menon has other interests and is not really concerned about assisting positivist research to get "a new lease on life" (Menon, 1993 p. 13).

More generally, citing misuse of a method to criticise the method itself is misleading. A moderate view would be that SST can be a useful guide in association with other ways of expressing results, particularly when research is looking for guidelines, or for the most likely profitable line of enquiry. That is, when decisions need to be made.

Intended or not, the function of overstatement is to produce a non-helpful reaction from all but the most committed of disciples. Menon's case against SST is made more unpalatable than it need be, if reasoned debate is the aim, rather than confrontation.

Methodology and the Role of Theory

Menon's argument seems to be that an intention to use SST will tempt researchers to undertake poorly designed studies and/or to investigate relatively unimportant hypotheses. Of course some research is poorly designed, and the work of some researchers is below standard, but any suggestion that there is a causal link between intended use of SST and poor research is not sustainable on the evidence presented.

The more general argument here concerns the importance of theory. A theory is essential as the basis for all research whatever methods are adopted, and one hopes that the application of "informed judgement" is based on theoretical rigour as well as on experience. Even when investigative research is primarily concerned with the development of theory it still requires at least a rudimentary theory to begin with.

The "Ideal World" of Educational Research

Menon is accurate in noting the lack of replication of research in education. In an ideal world replication would be undertaken as a matter of course. However, the reason that replication is rare in education, quoted and approved by Menon, "... the aura of respectability, replicability and generalisability ... [of] research that is based on SST ... inhibits actual replications ..." (1993, p. 13) is suspect at best. Researchers acknowledge that replication of studies is desirable, but also acknowledge that replication requires less intellectual effort than breaking new ground.

By far the majority of research is done by graduate students, and we want our

doctoral and research masters students to demonstrate more than the ability to replicate the research of others. It is also necessary to consider realities imposed by the restricted availability of funding. While the situation remains that there is barely enough money available for research in education to undertake any studies at all, doctoral studies take precedence, other new knowledge studies come next in the funding queue, and replication studies receive what, if anything, is left.

Summary

Let me first touch on two areas where I can more unreservedly agree with Menon. First, research of any kind should be published according to quality, not according to significance of results. I do not know any journal editors in the 1990s who would dispute this.

Second, Menon (1993, pp. 14–15) argues that more information about results of quantitative research should be reported and suggests several types of reporting. As he accepts, some of these practices are now more often followed than previously. In particular, measures such as simple descriptive statistics and the more complex effect size measures are becoming common. Among other benefits, such as making it possible for a reader to re-analyse, in part, the data presented, more complete reporting of results recognises and reveals the complex nature of educational theory and practice, and consequently the need to use mulitvariate approaches to analysis, whether quantitative or qualitative. With the development of techniques such as path analysis (for an example related to mathematics teaching and learning see Bourke, 1984), structural equation modelling (e.g., Marsh, 1989) and multilevel analyses (e.g., Cheung, Keeves, Sellin, & Tsoi, 1990), quantitative exploration of educational phenomena has become much more situational and realistic, although at a considerable cost in complexity.

However, at the extreme (and this is where Menon's article takes us), the policy cited with approbation by Menon would restrict the exposure to different genre of research in our professional journals. In the 1970s Carver was as guilty as editors of the 1960s journals he attacked in seeking to restrict publication of what could be termed as ideologically incorrect research. Belief in the one true way of conducting research is, unfortunately, not dead in the 1990s and, in Australia at least, the expression of this belief is not being espoused by adherents of the empirical tradition. Apparently Menon also objects to the continued existence and publication of research using SST (1993, pp. 13–14).

Researchers interested in reading reports of studies in mathematics education exploiting the widest variety of methods can only trust that the editor of this particular journal will not be employing criteria such as the use of SST as grounds for an editorial policy on what will not be accepted. This concern is related to the point made above about not publishing research where differences were not found to be significant. The exclusion of articles on methodological preference bases (i.e., to move away from quality and relevance as the criteria for publication in a journal) would be as unacceptable today as it was when Menon refers to it as existing in the past.

References

Bourke, S. F. (1984). *The teaching and learning of mathematics.*. Hawthorn, Vic: Australian Council for Educational Research.

Cheung, K. C., Keeves, J. P., Sellin, N., & Tsoi, S. C. (1990). The analysis of multilevel data in educational research: Studies of problems and their solutions. *International Journal of Educational Research, 14,* 215–319.

Marsh, H. W. (1989). Sex differences in the development of verbal and mathematics constructs: The high school and beyond study. *American Educational Research Journal, 26,* 191–225.

Menon, R. (1993). Statistical significance testing should be discontinued in mathematics education research. *Mathematics Education Research Journal, 5*(1), 4–18.

Author

Sid Bourke, Faculty of Education, The University of Newcastle, New South Wales 2308, Australia.