

Journal watch



Special retro edition

by Adam Jacobs

This edition of journal watch fails to bring you the latest developments from the medical literature, but rather looks back to times gone by, when there were nonetheless some fascinating papers published. One of the papers I shall write about here, from 1973, is apparently a true classic, but as I had never heard of it until recently I thought perhaps that other TWS readers might also be unfamiliar with it. So today I bring you 3 thought-provoking papers published between 1973 and 1980. They share a common theme of telling us some unsettling things about the way we perceive written (or even spoken) descriptions of research.

On to the first paper then, published in 1973 by Naftulin et al, which describes an experiment known as 'The Doctor Fox Lecture' [1]. The authors were all educationalists, who were interested in the effectiveness of teaching and how it is measured. Although it was common for educators to be assessed by having students complete satisfaction questionnaires (and indeed still is), Naftulin et al were concerned that student satisfaction with a specific teaching encounter is a poor measure of how much students actually learned, being more influenced by the charisma or popularity of the lecturer than by the effectiveness of the lecture content. They wondered whether not only students, but also professional educators, would be similarly influenced. They thus described the hypothesis of their study as follows:

"Given a sufficiently impressive lecture paradigm, an experienced group of educators participating in a new learning situation can feel satisfied that they have learned despite irrelevant, conflicting, and meaningless content conveyed by the lecturer."

To test this hypothesis, they embarked on what sounds like a tremendously fun piece of research. They recruited a professional actor, invented an impressive-looking CV for him that made him appear to be an expert on the application of mathematics to human behaviour, gave him the name of 'Dr. Myron L. Fox', and had him deliver a lecture to a group of professional educators from the fields of psychiatry, psychology, and social work, as part of a conference designed to help them become more effective educators of other health professionals. His lecture was on the topic of 'Mathematical Game Theory as Applied to Physician Education'. Naftulin et al coached him in how to deliver the lecture and handle the subsequent question and answer session "with an excessive use of double talk, neologisms, non sequiturs,

and contradictory statements." The lecture was videotaped and subsequently shown to 2 similar groups of professional educators.

Despite the nonsensical nature of the talk, most of the educators gave favourable responses in their feedback questionnaires, and none of them realised that the lecture was not genuine. Although there are some limitations to this study, particularly the lack of a control lecturer (we don't know how the group of educators would have rated a genuine lecturer), the authors' conclusions that student satisfaction scores are influenced more by style than by substance seems perfectly plausible, a worrying thought as I write this on UK general election day.

As an aside, I have my suspicions that I may have sat through a repeat of this experiment earlier this year when I listened to an utterly incomprehensible lecture on the philosophy of research ethics review, as regular readers of my blog will know [2].

Another paper, published in 1980 by Armstrong, followed on from the Dr Fox experiment by investigating whether the same effect was applicable to the written word. Armstrong summarised the 'Dr Fox hypothesis' as follows: "An unintelligible communication from a legitimate source in the recipients' area of expertise will increase the recipient's rating of the author's competence." He then went on to point out that the purpose of scientific writing should, looked at rationally, be to communicate knowledge, but that this conflicts with a desire for scientific writing to increase the writer's prestige if unintelligible writing is rated by readers as being more impressive.

He investigated this by 2 separate methods. First, he took a sample of 10 management journals, and rated each one for readability using the Flesch reading ease test. He then determined the prestige of each journal by asking a sample of 20 academics in the field to give each journal a rating. He found a significant correlation between the complexity of writing and the prestige of the journal.

Realising that one possible explanation for the results was that more prestigious journals handle more complex topics, requiring more complex language, Armstrong then went on to test his hypothesis with a further experiment. He took the conclusions sections from papers in 4 management journals, and rewrote them to alter their readability scores without affecting the content. He produced both simplified

and more complex versions for each passage. He then asked academics to rate the competence of the research described in the articles. Based on 32 responses, he found that passages written in a simplified style were rated of significantly lower competence than the others, despite the way in which he had controlled for the nature of the research described.

This makes depressing reading for medical writers. As medical writers, we are trained to write in a style that is straightforward and easy to read. Although this is a good way to communicate information, it is actually not a good way to impress your readership, who may be more impressed if you write in a more complex and less easily understood style. Perhaps this explains a phenomenon that I'm sure most medical writers have experienced when a document comes back with a client's edits, many of which seem to make the document less easy to read. In the 30 years since Armstrong's paper was published, I doubt that very much has changed.

My own perspective on the above 2 studies is that most people, particularly if they are supposed to be knowledgeable in a subject, don't like to admit if they have failed to understand something about that subject. There is probably a tendency for people to assume that the reason why they have not understood something is a function of their own poor understanding, rather than the poor communication of the person who delivered the information. The world would be a better place if we were all not afraid to say quite clearly when we don't understand something.

My last paper, published by Mahoney in 1977 (round about the same time as the picture of me at the top of this article was taken), looks at the way cognitive biases can affect the peer review process [4]. I first came across this paper a few years ago when working on a paper on cognitive biases in medicine [5]. I thought it was fascinating then, and I still do.

Mahoney was concerned about the extent to which peer review of academic journals might be affected by a phenomenon that psychologists describe as confirmatory bias: the tendency for humans to welcome experiences that support their pre-existing beliefs, and to be suspicious of anything that contradicts them. To investigate this, Mahoney wrote 5 different versions of a manuscript describing a fictitious experiment on the effects of extrinsic reinforcement on intrinsic interest, which was a controversial topic among psychologists at the time. He then sent the manuscript to peer reviewers, who were unaware that they were taking part in an experiment and asked to rate the manuscript on various aspects of its quality. The reviewers' perspective on the controversial topic was inferred by their association with a journal which had taken a clear line on the controversy. Some versions of the manuscript gave results that were consistent with the reviewers' presumed pre-existing ideas, and some gave results that contradicted those

ideas. The introduction and methods sections of the manuscript were identical in all cases.

The reviewers' ratings of the manuscript were significantly affected by the content of the results section. Crucially, there were significant differences in the reviewers' ratings of the methods sections, despite identical methods in all versions of the manuscript. As you have probably guessed by now, reviewers rated the methods as being of better quality when the results supported their pre-existing beliefs. Not only that, but the reviewers were more likely to recommend such manuscripts for acceptance in the journal.

Mahoney's paper shows very clearly that peer reviewers of papers are just as susceptible to cognitive biases as anyone else, a fact that probably doesn't come as a great surprise to most medical writers. He concludes "Without further scrutiny of the purposes and processes of peer review, we are left with little to defend it other than tradition." There may have been plenty of further scrutiny of peer review in the intervening 33 years, but I doubt that there is much more to defend it despite that further scrutiny. It is worrying that after all that time since the bias inherent in peer review was demonstrated so clearly, it still forms the cornerstone of scientific publishing.

Adam Jacobs
Dianthus Medical Limited
ajacobs@dianthus.co.uk
www.dianthus.co.uk

References:

1. Naftulin DH, Ware JE, Donnelly FA. The Doctor Fox Lecture: A Paradigm of Educational Seduction. *J Med Educ* 1973;48(7):630-635.
2. Jacobs A. Mission Impenetrable. <http://dianthus.co.uk/mission-impenetrable>¹
3. Armstrong JS. Unintelligible Management Research and Academic Prestige. *Interfaces* 1980;10(2):80-86.
4. Mahoney MJ. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognit Ther Res* 1977;1:161-175.
5. Klein J. Five pitfalls in decisions about diagnosis and prescribing. *BMJ*. 2005;330:781-783

1 In keeping with the retro theme of this article, I have tried to use 1970s typographic and bibliographic conventions, although I must confess I'm struggling to remember what the convention for referring to a blog entry was back then.