Transgressing the boundaries of science: Glazer, scepticism, and Emily’s experiment

Thomas Cox RN MS MSW
Doctoral Candidate, School of Nursing, Virginia Commonwealth University, Richmond, VA, USA

Introduction

Recently, I read Allmark’s (2003) article, ‘Popper and Nursing Theory’, in which he refers to Glazer’s ‘Post-modernism and TT’ (Glazer, 2001) and realized that the Journal of the American Medical Association (JAMA) article (Rosa et al., 1998), and Glazer’s use of it, should be re-addressed in this forum. I hope to demonstrate that the JAMA article is deeply flawed and best used as an exemplar of the misuse of science. Hence, Glazer’s (2000, 2001, 2002) use of the JAMA article makes me wonder why we nurses care what Glazer thinks.

The JAMA article, cited by Glazer, details Emily Rosa’s testing of therapeutic touch practitioners (TTPs). I will address five statistical issues that suggest obvious flaws in the JAMA authors’ research design, data analysis, and conclusions. Anti-TT sceptics failed to note these five flaws and failed to challenge the JAMA conclusions. However, sceptics paid a high price to secure a small debating advantage by abandoning rational scepticism and the scientific method when using this article.

During the 1997–98 academic year, as a first-year nursing student at a small, parochial school, I found it difficult to distinguish nursing from medicine. I began to doubt my decision to enter nursing and questioned whether I should have chosen medicine instead. I valued scientific medicine but I sought something more, something that both embraced and transcended medical science and would address the spiritual, empathic, and caring dimensions I expected in nursing. Between my first two semesters, I took seminars on shamanism, massage, and spiritual healing at the Esalen Institute and left Esalen knowing what I was missing.

During my second semester, I presented on Martha E. Rogers. As I read Rogers’ book (Rogers, 1970), I felt hopeful again. Rogers seemed to think as I did, that nursing is not just about traditional science, but transcends science. Nursing is about remaining with patients when science and medicine have proved insufficient and long after other health professionals leave. I do not believe Rogers ever abandons science despite what some sceptics think (Glazer, 2000; Raskin, 2000). Instead, I believe Rogers was warning people that over-reliance on traditional, cookbook science, borrowed science from other disciplines, and insufficient curiosity would never create and sustain a unique science of nursing.

I joined New York University (NYU)’s Martha E. Rogers listserver (I now co-moderate the Martha_E_Rogers listserver on YahooGroups with Dr Fran Biley), and despite my first-year status and small-school affiliation, I was on equal footing with other members. There, fresh ideas and reasoned arguments seemed to count, a profound difference compared with my experiences at school. Absent my Rogerian friends, I would have dropped out of school. On 1 April 1998, someone mentioned the JAMA article on the NYU listserver. I earned a master’s degree in applied mathematics and statistics before entering nursing school, and after reading the JAMA article

Correspondence: Thomas Cox, 1711 55th Terrace, Gainesville, FL 32605, USA. Tel.: 804 677 4140; e-mail: tcox@hsc.vcu.edu
and noting several glaring problems, detailed below, I honestly thought the article was intended as an April Fool’s joke.

By 2000, no one had published a refutation of the JAMA article, although I had reviewed it in two presentations (Cox, 1999, 2000). I subsequently published a small piece (Hallman & Cox, 2000) and prepared a full article (Cox, 2003). Approaching the article sceptically, as Glazer and anti-TT sceptics should have, I posed the question: ‘Why did the TTPs perform so poorly?’ First, I looked to the conclusions and calculations thinking they might be wrong and tried to verify them, in the tradition of rational scepticism and the scientific method. This approach was particularly important in view of the authors’ clear anti-TT biases. Following this, I considered alternatives to the authors’ analyses and explanations. I also started participating on a sceptic-run listserver on Yahoogroups where I shared my concerns and my critique of the article. Sceptics uniformly refused my oft-repeated suggestions to independently replicate my re-analysis rather than accept the article or my critique uncritically.

The major assumptions and conclusions of the JAMA authors that TTPs failed a reasonable test are incorrect. This is no ‘old science’ vs. ‘new age’ paradigm clash. The JAMA authors, Glazer, and I all share the same paradigm, but they abandoned it when they uncritically published and endorsed the JAMA article. I believe in science, the scientific method, critical thinking, and rational scepticism. The JAMA authors commit the following errors: repeated misuse of or failure to correctly interpret and report the results of the most basic statistical tests; misleading data presentations; and erroneous calculations. Every error, except a very critical counting error, favours the authors’ biases and these errors are obvious.

While the statistical errors are severe, there are ethical concerns as well, including: problems of informed consent, the researchers’ failures to protect their research subjects from harm, and discrepancies between the authors’ published protocol and what they actually did. As a traditional scientist, I find this level of error and self-contradiction unacceptable in a journal article. I cannot understand why self-described sceptics accepted and endorsed this article (Raskin, 2000; Glazer, 2001). At best, the JAMA authors’ most significant conclusions are incorrect; at worst, they may constitute scientific misconduct. The JAMA editors’ failure to withdraw this article confounds me and harms unwary readers who continue to cite it.

**Critique of the JAMA article**

Are the JAMA authors’ conclusions consistent with their data? I (Cox, 2003) detail serious statistical problems in the JAMA article, summarized below. The conclusion that TTPs scored ‘worse’ than chance, is incorrect. The authors’ data show that TTPs performed better with one hand, contradicting another of the authors’ conclusions. The authors’ stated alpha and power values are incorrect and deviate from accepted biomedical research standards. The most glaring single error is the publication of two values for the number of correct answers in the same 280 trials (Rosa et al., 1998; Barrett, 2000).

The JAMA authors contend that the TTPs’ performance was: ‘… close to what would be expected for random guessing’ (Rosa et al., 1998, p. 1009). The two-tailed Binomial test is the correct approach to use in testing a random guessing hypothesis. True random guessers should score between 124 and 156 correct answers in 280 trials, in about 95% of such experiments. Higher or lower scores, such as the authors’ two reported values of 123 (122) correct, require rejection of a random guessing hypothesis because the a priori probability of both these outcomes is lower than 0.05, the standard value for alpha in biomedical research (\(P_{123|280} = 0.0484; P_{122|280} = 0.0362\)). The authors’ failure to reject the random guessing hypothesis is puzzling because the binomial test of random guessing is one of the most basic statistical tests covered in introductory statistics books. As well, hypothesis testing of this data results in a clear true or false outcome. ‘Close to’ has no meaningful statistical interpretation for this data.

A second, extremely disturbing problem, alluded to above, is that the JAMA article states 123 successes in 280 trials. Nevertheless, Barrett’s (2000) website states there were 122 successes and has not been changed despite this authors’ repeated questioning of
this value. This suggests that at least one of the JAMA authors believes that asserting two different numbers ‘correct’ in the same 280 trials is acceptable, despite the fact that these two statements violate one of the most fundamental rules of logic, the law of contradiction. Postmodernists may embrace contradiction, but do the JAMA authors and does Glazer? The error in the main conclusion and the counting discrepancy are extremely serious errors for an article published in a top-tier medical journal and hailed, by sceptics, as proof that TT is unscientific.

The authors’ problems include another incorrectly analysed and reported conclusion. They assert that there was no statistically significant difference in detection rates by hand for the TTPs in their first round of trials. In fact, they went further, disparaging their subjects for suggesting that such an effect existed (Rosa et al., 1998). There is no TT literature suggesting such an effect, making the authors’ case for performing this analysis extremely weak. The data presented by the authors suggest that when tested for a hand effect the TTPs answered correctly 27 out of 72 times with right hands and 43 out of 78 times with left hands. I (Cox, 2003) re-analysed this JAMA hand data and reported the appropriate conclusions using both Fisher’s Exact Test (P = 0.0344) and the chi-square test (P = 0.0306). Both these tests are appropriate analyses of the authors’ contingency table data and both require rejection of the authors’ hypothesis: no difference in performance by hand. For the third time, the authors’ statements and data are irreconcilable. Glazer, had she critically reviewed this article, should have noted such obvious problems.

However, things get worse for the JAMA authors and by extension, for Glazer. In fairness to Glazer, the next two problems require a deeper grasp of statistical reasoning. I (Cox, 2003) recalculated alpha and power values for Emily’s experiment and found that the authors made severe statistical design errors. They calculated and reported alpha and power values for a single phase, 10-trial test. However, according to their protocol, TTPs fail unless they pass two phases in which they get at least eight of 10 correct answers. Using two phases changes the experiment and the alpha and power values. Many sceptics do not understand this and they often ask for multiple testing without correcting alpha and power values. The JAMA authors report the correct alpha for a single phase as 0.055, but the correct alpha value for a two-phase test is actually 0.003025. This alpha value means that Emily’s experiment will not incorrectly validate a randomly guessing TTP except three times in 1000, a protection level inappropriate in a fourth grader’s science fair project and without parallel in standard biomedical research.

The last serious error occurred when the authors calculated the power for Emily’s experiment, leading to a misstatement of the likelihood that skilled TTPs would pass Emily’s experiment. Power refers to the probability that a TTP skilled at a prespecified level, at tasks identical to those tested, will pass the test. In fact, TTPs make no claims about their ability to detect hands so, appealing as it may be, the test may not apply to routine TT practice at all. The JAMA authors failed to adjust their power calculations for the two phases required, resulting in a clear bias against TTPs in Emily’s experiment compared to generally accepted biomedical research standards which call for a power of at least 0.80. The authors suggested that TTPs with moderate skills such as two out of three (three out of four) correct answers should pass Emily’s experiment. However, the true power of Emily’s experiment, at the authors’ specified skill levels, is less than 0.10 (0.28). This bears repeating: at the authors’ specified skill levels there is a 90% (72%) chance that TTPs will fail the test. There is simply no justification for these values or their incorrect calculation in an article published in a journal of the status and regard of JAMA.

Conclusion

I have detailed five serious statistical errors in the JAMA paper. Any reader willing to consult an introductory statistics text can validate this re-analysis for himself or herself. There is no paradigmatic conflict. Glazer and I probably agree that things are either right or wrong. Many sceptics, like Glazer, uncritically endorsed this flawed article and abandoned science and critical thinking in the process.

Glazer’s ideas might be worthy independently of her failure to critically assess the evidence she used.
However, as I suspect would be true for Glazer, I see no reason to forgive her lack of critical judgement in endorsing the JAMA article. I believe she may have some good ideas, but her failure to critically assess her evidence makes me wonder if she really has anything to offer science, nursing science, or nursing philosophy. We, in nursing, already have ongoing discussions on epistemology and methodology. Glazer superficially echoes what we already know, and then dramatically, albeit unintentionally, proves that when people allow prior beliefs to cloud their judgement, they forget to think critically. Failure to think critically when you think you are right appears to occur independent of paradigm.

Nobody, not sceptics, not Glazer, not physicists, not nursing scientists of either positivist or postmodernist persuasion, have a lock on knowledge or a philosopher’s stone. Anyone at all who relies on methodology and forgets the value of self-doubt could fall for the unsupported conclusions and assertions in the JAMA article. I believe that the errors made by Glazer are precisely the errors Rogers warned her students and colleagues to avoid. Rogers knew nursing science and nursing research must attend to the heart, mind, and soul of nurses, nursing, and nursing clients, not just one and not with a prepackaged set of tools. I believe Rogers knew the agenda needed science and rigor, but also heart, soul, and faith in the unexpected, because nursing is incomplete without all of these. I believe Rogers knew that no paradigms, computers, or research methods are magic bullets for knowledge development. Instead, we must use all methods wisely and appraise them critically. We in nursing should continue in our disciplinary dialogues, learning together from our best efforts and worst mistakes. I do not believe we need help sorting this out from someone who, upon entering our dialogue, should have exposed the JAMA article, not endorsed it.

References