Forensic Science International 277 (2017) 66–76

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Review Article

Peer review in forensic science

Kaye N. Ballantyne[a,b,*], Gary Edmond[c], Bryan Found[a,c]

[a] Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod Victoria, Australia
[b] School of Psychology and Public Health, La Trobe University, Bundoora, Victoria, Australia
[c] Program in Expertise, Evidence and Law, Faculty of Law, University of New South Wales, Kensington 2052, Australia

ARTICLE INFO

Article history:
Received 24 November 2016
Received in revised form 3 April 2017
Accepted 17 May 2017
Available online 25 May 2017

Keywords:
Peer review
Verification
Cognitive factors
Expert evidence
Report

ABSTRACT

Peer review features prominently in the forensic sciences. Drawing on recent research and studies, this article examines different types of peer review, specifically: editorial peer review; peer review by the scientific community; technical and administrative review; and verification (and replication). The article reviews the different meanings of these quite disparate activities and their utility in relation to enhancing performance and reducing error. It explains how forensic practitioners should approach and use peer review, as well as how it should be described in expert reports and oral testimony. While peer review has considerable potential, and is a key component of modern quality management systems, its actual value in most forensic science settings has yet to be determined. In consequence, forensic practitioners should reflect on why they use specific review procedures and endeavour to make their actual practices and their potential value transparent to consumers; whether investigators, lawyers, jurors or judges. Claims that review increases the validity of a scientific technique or accuracy of opinions within a particular case should be avoided until empirical evidence is available to support such assertions.

Contents

* Corresponding author at: Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod Victoria, Australia.
E-mail addresses: kaye.ballantyne@police.vic.gov.au (K.N. Ballantyne), g.edmond@unsw.edu.au (G. Edmond).

http://dx.doi.org/10.1016/j.forsciint.2017.05.020
0379-0738/

1. Introduction

Peer review is one of the central components of the scientific framework underpinning the publication process in journals, the awarding of grants and honours, and promotion of academics. It has long been held up as the premier approach to ensure the validity of methods and conclusions, to detect errors and fraud, and to improve the quality of learned papers [105]. Courts have used peer review as an indicator of 'good science' and general acceptance within the relevant communities of experts, with landmark rulings such as Daubert and Kumho deeming peer review as an important factor in determining whether a scientific method can be accepted as valid [1,2]. The forensic sciences have universally adopted peer review, most conspicuously verification, as an essential part of quality management and error mitigation systems. Accrediting bodies have mandated case file review as part of standard quality control procedures, and professional societies have recommended the use of verification or review to ensure the soundness of conclusions drawn, and as a way of reducing error rates inherent in subjective methods.

Notwithstanding its long and widespread use, the value of peer review is frequently exaggerated, an outcome that may be the result of the variety of meanings attributed to the term. There is little evidence of the effectiveness of either peer review or verification. Among lawyers and forensic scientists there appears to be limited awareness of concerns about the ability of peer review, in any of its guises, to ensure methodological soundness or detect error and fraud. Indeed, it is seldom appreciated that in many of the high profile cases of known erroneous identifications or miscarriages of justice, peer review and verification failed to detect the error (e.g. [3–5]). Likewise, independent reviews of problematic laboratories and units within the United States have indicated that technical review procedures were inadequate, non-existent or completely undocumented, performed long after the report was issued, or that case file contents were so incomplete as to make a thorough review impossible [6–8].

There has been concern among many forensic scientists that the error rates cited in the PCAST report [9] are inaccurate and unrepresentative of true case work error rates, due to an absence of verification and review procedures in black box studies. For example, the OSAC Friction Ridge Subcommittee Response to PCAST indicated that the as the quoted black box studies do not contain *any* verification, the error rate "is expected to be lower, perhaps to a substantial degree, than those values reflected by the PCAST" [10]. Likewise, the Association of Firearm and Tool Mark Examiners (AFTE) regard the recommendation that court testimony refer to error rates from a single study performed on firearm examination as "irresponsible and inaccurate", in part due to the lack of technical and quality review processes in this study [11]. However, the claim that verification or review will lower error cannot be substantiated with empirical data in most disciplines, where error rates and distributions are unknown.

The risk of exaggerating the effectiveness of the various forms of peer review encountered across the forensic sciences is serious, and overt reliance on the practice to prevent errors may not be achieving desired aims. We introduce a taxonomy of review and verification processes, applicable to both scientific publications and forensic opinions. Our aim is to encourage transparency in order to facilitate more reliable estimation of the ability of peer review to contribute to the accuracy of evidence produced by forensic practitioners.

1.1. What is peer review?

The term 'peer review' is used to describe a range of different practices, used for a variety of purposes. Scientific articles are commonly subjected to editorial (pre-publication) review, where works are scrutinised by knowledgeable peers from a relevant field. Forensic reports and statements are checked through a process of technical and administrative review, ostensibly to ensure the accuracy and completeness of the opinion and associated documentation. Verification, within the forensic sciences, might involve replication (i.e. independent analysis or re-analysis) or just a review of the original examiners' analysis and opinion(s) to confirm the result and prevent erroneous opinions being reported. While all are collectively referred to as peer review, they have very different aims, involve different methods of review and the evidence of effectiveness varies. Below we describe the broad range of peer review applications in the context of both mainstream academic science and within the forensic sciences, followed by an examination of the evidence for the effectiveness of the forensically relevant review types in relation to the aims of the process.

1.1.1. Editorial peer review

Within academic (and some commercial) scientific domains, peer review is primarily a checking process, where two or three individuals, knowledgeable in the field, scrutinise papers to determine if the methodology is sound and applied in an appropriate manner, if the data produced has been correctly analysed with suitable statistical tests, and if the conclusions and recommendations drawn are appropriate to the breadth and depth of the study [12,13]. In most disciplines, reviewers do not, and cannot, replicate experimental methods or data—they must use their professional and scientific expertise to determine if the documented experimental design, methods, results and conclusions *appear* valid [12,13].

In the majority of cases, where review is for publication (or the award of research grants) reviewers spend less than 10 h reviewing submissions, with a median of 6 h across all disciplines [14]. Whilst reviewers scrutinise technical attributes of the research, as well as scientific quality, clarity of presentation and ethical validity [12], the review process does not conclusively authenticate or endorse the validity of the particular methods and conclusions. Instead, editorial peer-review, beginning with the Royal Society of Edinburgh in 1731 [13], was intended to assist editors in the selection of manuscripts for publication, by distributing material to "those members who are most versed in these matters". From the start, the ultimate responsibility for the integrity of the article lay with the author:

> "Responsibility concerning the truth of facts, the soundness of reasoning, in the accuracy of calculations is wholly disclaimed: and must rest alone, on the knowledge, judgement, or ability of the authors who have respectfully furnished such communications" [13].

Despite the early start to editorial peer-review, the practice was not formalised until the mid-20th century, with Science and The