Open science as a better gatekeeper for science and society: a perspective from neurolaw

Chuan-Peng Hu, Xiaoming Jiang, Ricky Jeffrey and Xi-Nian Zuo

Citation: Science Bulletin 63, 1529 (2018); doi: 10.1016/j.scib.2018.11.015

View online: http://engine.scichina.com/doi/10.1016/j.scib.2018.11.015

View Table of Contents: http://engine.scichina.com/publisher/scp/journal/SB/63/23

Published by the Science China Press

Articles you may be interested in

Open data for better science
National Science Review 5, 593 (2018);

Evolution of computers and simulations: from science and technology to the foundations of society
SCIENCE CHINA Chemistry 57, 1317 (2014);

Addressing the issue of fog and haze: A promising perspective from meteorological science and technology
SCIENCE CHINA Earth Sciences 57, 1 (2014);

Development of Science China Chemistry during 2008–2012: From the perspective of Special Issues/Topics
SCIENCE CHINA Chemistry 55, 2617 (2012);

The China power grid: a network science perspective
National Science Review 1, 368 (2014);
Commentary
Open science as a better gatekeeper for science and society: a perspective from neurolaw
Chuan-Peng Hu a,*, Xiaoming Jiang b, Ricky Jeffrey c, Xi-Nian Zuo d,*

aDeutsches Resilienz Zentrum (DRZ) & Neuroimaging Center (NIC), University Medical Center of the Johannes Gutenberg University, Mainz 55131, Germany
bDepartment of Psychology, Tongji University, Shanghai 200092, China
cCAS Key Laboratory of Behavioral Science, Institute of Psychology, the Chinese Academy of Sciences, Beijing 100101, China

In 2011, a court near Milan, Italy, reduced the sentence of a convicted murderer from life to 20 years in jail, after considering results from neuroimaging and genetic tests. These tests had been used to convince the judge that the trial's female defendant, Stefania Albertani, was suffering from partial mental illness (http://blogs.nature.com/news/2011/09/italian_court_reduces_murder_s.html). The structural MRI images showed that Albertani's grey matter volume was different from the average volume of a randomly-selected group – 10 healthy women with no history of mental or neurological disorders – in brain areas including the anterior cingulate gyrus (ACC) and insula. The expert testimony concluded that “these alterations (of brain structures) have to be considered in causal relation to the psychiatric symptomatology of the murderer.”

This is one of the many cases where structural or functional brain imaging results have been cited as evidence in the courtroom, alongside other evidence, in order to support claims that a defendant was mentally ill [1]. Which findings from cognitive neuroscience can be used as evidence in the courtroom (in the form of expert testimony) is thus a central question of the interdisciplinary neuroscience and the law [2]. Judges must assess the scientific validity and reliability of the cited scientific evidence in order to decide its admissibility. This constitutes a significant challenge, since judges may have limited training in science.

The recent reproducibility crisis in science, affecting medical science, neuroscience, and behavioral science among other fields, has made this situation worse [3]. (For a comprehensive reference list, please see https://osf.io/xunjji/) Studies suggest that the false positive rate in the psychological and neuroscience literature is surprisingly high [4]. The legal system, which relies on the scientific community to provide scientific evidence (Fig. 1), must thus now consider how to deal with this low reproducibility rate in published results.

Irreproducible science compromises the admissibility standards for scientific evidence. The Daubert standard from the US legal system offers one example. It requires judges to evaluate the expert evidence (scientific evidence included) when deciding whether this evidence is admissible, and has been intensively studied both in the United States and in other countries.

The standard requires that judges evaluate the validity of scientific evidence based on four criteria: (1) Is the opinion testable and has it been tested? (2) Is the error rate associated with the technique or opinion acceptable? (3) Has the basis for the opinion survived the peer review process and been published? (4) Is the technique or procedure for generating the opinion generally accepted among scientists in that field?

The Daubert standard suggests that, even though the judges are required to assess the scientific evidence themselves, they still rely heavily on input from the scientific community. Standards (iii) and (iv) require that the theory or technique behind the relevant scientific evidence has been approved by scientists. Thus, the scientific community acts as a gatekeeper for science in the first stage (Fig. 1).

According to the Daubert standard, if scientific results are flawed and yet nonetheless published, judges may have difficulty in excluding any evidence based on such results. To illustrate this impact of published but irreproducible results on the admissibility of scientific evidence, we can consider the Albertani case mentioned earlier. In this case, as in many others [1], the observation of abnormal brain structures in a single MRI scan was cited as evidence in expert testimony to infer the mental state of the offender.

Before analyzing how irreproducible research compromises the Daubert standard, we need to first assume that claims from experts are testable and have been tested, i.e. that the relationship between the ability to control one's behavior and the brain structure underlying this ability can be tested. At first glance, many published studies seem to report cognitive neuroscience findings that are linked to defendants' self-control. For example, behavioral flexibility, the ability to control one's attention, is a cognitive process frequently studied in psychology and cognitive neuroscience, and is supposedly related to self-control. Studies found that variation in behavioral flexibility is correlated with the thickness of grey matter in the lateral prefrontal cortex and ACC [5]. However, the concepts of self-control in law and in psychology and cognitive neuroscience are not always the same thing: while self-control in

* Corresponding authors.

E-mail addresses: hcp4715@gmail.com (C.-P. Hu), zuoxn@psych.ac.cn (X.-N. Zuo).

https://doi.org/10.1016/j.scib.2018.11.015
2018 Science China Press. Published by Elsevier B.V. and Science China Press. All rights reserved.

the law is associated directly with behaviors, criminal behaviors in particular, self-control in psychological and cognitive neuroscience studies is operationalized with some specific task that might not be directly associated with behaviors. In other words, when considering expert opinion on offenders’ ability to exert self-control, at worst this opinion may not even be testable, and at best it has not yet been tested.

Further, even if we assume that the operationalized tasks in psychology and neuroscience are indeed related to self-control ability in the legal sense, as many lawyers assume when citing brain imaging results, irreproducible results can still compromise the *Daubert* standard.

First, it is difficult to estimate the error rate based only on published papers, hence compromising the second standard of *Daubert*. The error rate indicated in the *Daubert* standard is in most cases interpreted as statistical error and method error. While the former error mainly refers to the differences between the predicted and the real values, estimated by measure of uncertainty (e.g. standard deviation), the latter error is more inherent to the methods itself (i.e. the sensitivity, meaning the proportion of true positives that are correctly identified by the method, and specificity, meaning the proportion of true negatives that are correctly identified by the method) [6]. The sensitivity of a brain imaging technique in detecting the relationship between brain structures and behavior might be low if the relationship reported in one brain imaging study cannot be reproduced in a direct replication. Indeed, the above mentioned study about the relationship between behavioral flexibility and the volume thickness of grey matter in the lateral prefrontal cortex and ACC [5] has not been replicated successfully [7]. Further, it would be impossible to estimate the sensitivity and specificity of using the brain imaging technique to detect brain-behavior relationship if the publicly reported results have been selected by publication bias. The existence of publication bias means that the literature does not reflect positive and negative results in an objective way. For example, on the relationship between brain structure and self-control, if only positive results have been published, we would not be able to evaluate the reliability and robustness of these results [8].

Second, the peer review and publication standard of *Daubert* is compromised by publication bias, which has inflated the false positive rate in the cognitive neuroscience literature [8]. Recent studies have revealed that the flexibility researchers have when selecting their own analytical approach to generate results is of a high level [9]. The combination of researchers’ flexibility and publication bias may have resulted in inflation of false positives in the cognitive neuroscience literature. In this situation, the published results are untrustworthy, and therefore compromise the peer review and publication standard.

Third, in the field of cognitive neuroscience, especially for neuroimaging studies, there is no generally accepted standard for methods. Almost every paper reports having used a different workflow to process the data [9]. The standardization of data processing has not received sufficient recognition, though a recent publication calls a consensus to be developed in data processing routine [10]. Given this proliferation of different research procedures, it is difficult to say that any procedure referenced in a legal case is generally accepted among the community of cognitive neuroscientists or not.

In sum, the reproducibility problems in cognitive neuroscience undermine the legal gatekeeping process (the *Daubert* standards). The situation becomes even worse if both the experts (clinical psychologists or neuroscientists) and the judges lack necessary training in statistics.

The open science movement can benefit the gatekeeper system. The scientific community has taken actions to address the credibility crisis caused by the reproducibility issue. The most prominent approach is the open science movement, which is characterized by openness, transparency, and reproducibility in scientific research practice. Direct replications are encouraged [3]; pre-registration before collecting data is becoming a new norm, e.g. [11]; open data, materials and code are required; new inferential statistical techniques are introduced and new tools developed, e.g. [12]; and standard protocols to analyse data and report results are proposed, e.g. [10]. In sum, the *open science movement* has started to clean out the influential but irreproducible studies, and increase the transparency, rigor and reproducibility of new studies, so as to enhance science’s ability to describe, explain and predict.

The benefit of the open science movement to science itself has been discussed at great length recently, and is beyond the scope of the current discussion (see https://osf.io/xunjy/ for a comprehensive reference list). Below we will focus on how the open science movement can benefit the gatekeeping process for admitting scientific evidence into the courtroom.

First, the open science movement may help to build a better system for evaluating error rates. In the current publishing culture in the field of cognitive neuroscience, as well as across science generally, direct replications are discouraged. As a result, novel findings are seldom replicated, which makes it hard to estimate the error rate. The open science movement has addressed this issue by emphasizing the value of direct replication [3]. Thus, open science allows the development of a more cumulative science, and helps us evaluate the error rate of a particular technique or method. In addition, open data itself provides greater opportunity to evaluate methods and techniques. For example, the Consortium for Reliability and Reproducibility (CoRR) has developed a range of useful resources to assess the reliability and validity of many methods used in neuroimaging research [13].

Second, the open science movement may help to build a better peer review system, which not only includes the usual pre-publication review but also a post-publication review [14]. Previously, the scientific community relied only on pre-publication peer-review, which cannot fully evaluate the reliability and validity of a study since only the results are visible to reviewers. The open science movement not only enhances the pre-publication review, but also emphasizes the value of post-publication review. The pre-publication review is enhanced by pre-registration, open materials, scripts
and data because in this case the evaluation of a study is based on the entire research process. The post-publication review, which is encouraged by the open science movement, also become valuable if the open scripts and open data are made public along with the publication (e.g. http://www.russpoldrack.org/2013/04/how-well-can-we-predict-future-criminal.html).

Third, the open science movement helps to build generally accepted standards which can then lessen researcher flexibility. Currently, nearly every neuroimaging study uses its own workflow and its own standard for analyzing the neuroimaging data [9], and this granting of full flexibility to researchers inflates the false positive rate. In the open science movement, a set of best practice is recommended [10], and the whole field is beginning to establish a generally accepted standard. These best practice standards may in time constitute a scientific routine that helps experts to better evaluate the reliability of scientific evidence in the courtroom.

In sum, the open science movement increases the rigor and reproducibility of science, and renders the whole research process more transparent to both researchers and the public. In this way, open science may help scientists to become better gatekeepers for science, which in turn will help judges to use the Daubert standard for better decision-making.

Based on the above analyses, we suggest that both cognitive neuroscientists and legal practitioners would benefit from welcoming the open science movement. The open science movement is not a simple standard, but “a set of beliefs, research practices, results, and policies that are organized around the central roles of transparency and verifiability in scientific practice” (https://twitter.com/mcxfrank/status/1044254887075147776), involving a dynamic, upward spiral process. We suggest that both cognitive neuroscientists and practitioners adjust their practices so that they are based on the best available standards.

For cognitive neuroscientists, it is not easy to follow the newly established standards, which require the preregistration of a study, complete reporting, and the opening of all data and scripts. Fortunately, many practical guides have been published to help researchers navigate through open science [12,15]. Professional organizations have also begun to develop standards to enhance the rigor of science, e.g. the Organization for Human Brain Mapping (OHBM) has created the Committee on Best Practices in Data Analysis and Sharing (COBIDAS; http://www.humanbrainmapping.org/cobidas) to guide analysis and sharing of MRI and M/EEG data; while the American Psychological Association has also published standards for reporting both quantitative and qualitative studies. These practical primers and standards provide researchers with a guide on how to do better as gatekeepers of science.

For legal practitioners, it is crucial to be aware of the current irreproducibility problem and its potential challenge to the credibility of scientific evidence. However, scientific journal articles, including those on reproducibility and open science, are often too technical and cannot be easily understood even by researchers, let alone legal practitioners. Therefore, to be better gatekeepers, it is important for legal practitioners to collaborate closely with researchers. In neurolaw, there is a tradition of neuroscientists working closely with legal scholars and legal practitioners. For example, in 2010, with the support of the MacArthur Foundation, a group of neuroscientists edited a guidebook, A Judge’s Guide to Neuroscience, which explained the basics of neuroimaging studies in plain language. It would be helpful if there were similar guides for judges and other legal practitioners about reproducibility and open science. If judges are made aware that published results can be false positive and irreproducible, and that researchers may, intentionally or not, p-hack for publication, they are likely to be more cautious when using scientific evidence from experts in their decision-making. This awareness, further raising the bar set by the Daubert standard, may prevent additional unreliable scientific evidence from entering the courtroom.

Conflict of interest

The authors declare that they have no conflict of interest.

References


Chuan-Peng Hu is a Post-doc researcher at the German Research Center (DRZ) and the Neuroimaging Center (NIC), University Medical Center of the Johannes Gutenberg University. His research interests include (1) neural basis of the self-enhancement effect and psychological resilience, (2) reproducibility of psychological and brain sciences, and (3) neurolaw and neuroethics.

Xi-Nian Zuo is a full research professor of psychology, director of the Research Center for Lifespan Development of Mind and Brain (CLMB) and Magnetic Resonance Imaging Research Center (MIRRORC) at the Institute of Psychology, Chinese Academy of Sciences. His research interests center around reliable human brain mapping. He founded both Consortium for Reliability and Reproducibility (CoRR) and Chinese Color Nest Project (CCNP), and currently serves the Organization for Human Brain Mapping (OHBM) as a council member and the program committee chairperson.