International Journal of Risk & Safety in Medicine 13 (2000) 109–116
IOS Press

Four editorial proposals to improve social sciences research and publication

Jay Joseph [a,*] and Steve Baldwin [b]

[a] *Berkeley, CA, USA*

[b] *School of Social Sciences, University of Teesside, Middlesbrough, UK*

Abstract. This paper discusses unresolved key problems in social sciences research and publishing. Researcher bias can lead to the publication of studies with highly questionable conclusions, which can impact adversely on important social topics (e.g., the influence of heredity on intelligence, criminality). Outright fraud occurs in science, and social sciences in particular [30]. To help counter the prevailing publication bias in research, four editorial suggestions for scientific reporting are proposed: (1) Before the work starts, investigators would be required to submit a written description of their research methodology. An international social sciences evaluation registry would be created to monitor research activity. (2) The Editor's decision about whether to publish a research study would be based on a paper that omitted the results and conclusions sections. These sections would be included in the published paper after an editorial decision was made. (3) With these appropriate scientific safeguards, researchers would also be required to provide raw data to other interested scientists, especially when the study is socially important, or is difficult to replicate. (4) Journal Editors who greenlight socially relevant studies would provide extra publishing space to people whose views differ significantly from the principal investigators. Bias and distortion cannot be completely eliminated from the scientific publication process. Adoption of these four proposals, however, would reduce scientific reporting problems such as data dredging and the dissemination of unsupported conclusions, and would also help to reduce fraud. Such adoption would eventually lead to the publication of improved scientific research in the social sciences.

1. Introduction

The conduct, reporting, and publication of social science research requires an urgent overhaul. Current major unresolved problems include: (1) the reluctance of journal Editors to publish non-significant results or replication studies, (2) the shoddy practice of researchers changing definitions and methods to find and then report results that correspond to their original hypotheses, beliefs, or vested interests, (3) the tendency to accept the investigator's conclusions, to the exclusion of alternative plausible explanations of the results, and (4) the presence of outright fraud.

Since the early 1970s, several authors have made impassioned calls to change the existing publication system in psychology and related social science fields (e.g., [16,29,31]). Minor problems in the publication process (e.g., inadvertent mistranslation, incorrect date citation) probably cause insignificant scientific damage. Major problems have developed, however, when ethical codes have been missing in science. For example, twin studies in psychiatry and psychology often have been cited to support the importance of genetic factors in psychological trait differences [21,23]. Many challenges have been made since the mid-1970s to Cyril Burt's claim to have administered IQ tests to 53 pairs of reared-apart monozygotic twins [18,26]. Without Kamin's diligence, major inconsistencies in Burt's data might not have been noticed. If Burt had faked his data better, there would have been fewer clues (i.e., "invariant

---

[*] Requests for reprints can be sent to: Jay Joseph, 2625 Alcatraz Avenue, #328, Berkeley, CA 94705, USA. E-mail: jayjoseph2@aol.com.

correlations") to be discovered. Although some authors have attempted to salvage Burt's damaged reputation (e.g., [14,25]), the contemporary debate has focused on whether inconsistencies in Burt's data resulted from purposeful fraud or "research carelessness". Irrespective of partisan factors, psychologists generally have agreed that the Burt papers on reared-apart twins are no longer worthy of scientific consideration. Rightly, the status of Burt's other work has been contested. Burt likely invented not merely a dataset, but also the existence of his researchers [32]. Many well-informed contemporary psychologists believe that *none* of Burt's findings can be trusted.

Burt is not the only well-known researcher alleged to have faked or fudged his or her data, and fraud is rife in social science research (for a review, see [1,5,10,15,30]). In the 2000s, innovative methods are urgently required to promote fraud elimination in scientific work. The aim of the current paper is to propose four innovative methods to prevent or reduce manipulation of data by researchers. Data manipulation can be purposeful or "accidental"; in either instance, science is distorted to produce conclusions consistent with pre-determined personal biases. Savan [39] noted that scientists:

> are not human encyclopedias or data processors who need only to find the right file to come up with the single true answer to our question. Like traditional priests, they are individuals with personal histories, childhood experiences, phobias, religious and political convictions, hopes, goals, desires, and ambitions. They are, like all of us, emotional beings, pursuing, usually with passion and even obsession, work to which they are intensely dedicated. Inevitably these human qualities color a scientist's work, and influence the field, line of research, methods, and ultimately the conclusions of his or her investigations (p. 21).

It is impossible (and undesirable) to remove all bias from social sciences research. Rather, such distortions should be acknowledged and then subjected to counter-controls. Four proposals are made to protect science and society from the false ideas and unwarranted conclusions of biased scientists:

(1) *Before initiating a study, researchers should be required to submit a written description of how they will obtain subjects, how they will define and measure the variables of interest, and how the comparison of groups will be performed.*

An internationally-based social science central register should be established to collect and store this information. The register would create a permanent public record of the intentions and methods of researchers *before* data collection and analysis. A *research register* is "a database of research studies, either planned, active, or completed (or any combination of these), usually oriented around a common feature of the studies such as subject matter, funding source, or design" [12, p. 72]. Simes [40] noted that an international registry of clinical trials in medicine would register "all clinical trials at their inception with the objectives and endpoints of each trial clearly stated" (p. 12).

Research registers are already used in some areas of medicine and should now be extended to the social sciences. Cooper, DeNeve and Charlton [11, p. 452] have called for the creation of "an international, standardized, computer accessible register of IRB [institutional review board]-approved research projects" for psychology research studies. Cooper and colleagues proposed the register to combat publication bias problems. Such problems have been defined as "a bias against studies with negative findings, on the part of authors, reviewers, and editors, which can lead to the systematic exclusion of negative findings from the published literature" [41, p. 147]. A bias against inclusion of negative findings can produce considerable problems for research synthesis (meta-analysis) studies. De facto, meta-analysis studies can only include *published* work. Such publication bias also creates a distorted impression of research that is included in literature reviews. Publication bias creates false impressions about the body of research on specific topics, as studies that fail to reject the null hypothesis are rarely published [2–4,37,42,43].

The most important purpose of an international social science register (ISSR) would be to record researchers' methods and definitions before the collection of data – a point made by Easterbrook:

> A further advantage [of a registry] is that through access to information on stated prior hypotheses, it would be possible to identify the use of multiple comparisons (that is whether positive results were based on prior hypotheses or posterior analyses). Such information could be helpful in the publication review process [13, p. 346].

The register of studies would aid fraud prevention and help ensure that evaluators analyzed results according to their research design specification. Precise design specification would be completed before data collection. Given the abuses in social science research, a monitoring system is required. Sometimes, as the research proceeds, changes are required in the design specification. In these instances, a statement would be sent to ISSR to explain the protocol revisions. Ideally, information about research design sent to ISSR would be published in the form of a brief abstract before the publication of the final paper.

Registration of studies at ISSR would reduce the unscientific (but common) practice of "data dredging", that is, the purposeful post-hoc analysis/manipulation of datasets to locate significant relationships that support the investigator's research hypothesis. Third party financial interest groups can benefit directly from such data manipulations. Adoption studies about the "genetics of criminality" and problem drinking illustrate the impact of data dredging. The social perception of causes of sanctioned and disapproved behavior has been unduly influenced by such dredging. In the Bohman [6] Swedish adoption study, for example, no significant correlation for registered criminality was found between adoptees and their biological parents. Bohman concluded there were no genetic determinants for criminality. Four years later, however, Bohman and colleagues [7] published a paper that reanalyzed results from the 1978 study. According to the authors, there was a genetic influence on the commission of property crime. The Bohman et al. [7] study is often cited in support of the alleged influence of genetic factors on criminal behavior. There is no evidence however that Bohman and colleagues determined their comparison criteria in advance. It is well known that in a large dataset, probability dictates that by chance some significant correlations will be found between variables that have no causal or associative relationship [32].

Another example of data dredging is in the paper that reanalyzed the Danish/American schizophrenia adoption study of Rosenthal and colleagues [9,36]. In the reanalysis by Haier, Rosenthal and Wender [17], MMPI scores of the adopted-away offspring of people diagnosed with schizophrenia were compared with scores of the adopted-away offspring of controls. The authors described several comparisons that did *not* differentiate between the two groups of offspring. Eventually, the investigators produced one comparison between variables that *did* differentiate between the groups. The authors (incorrectly) concluded that the study "continued to support the genetic hypothesis" [17, p. 175]. The Haier et al. paper is a clear example of unacceptable methods of data analysis. Nonetheless, the authors candidly discussed how each comparison failed to support their hypothesis, compelling them to analyze their data differently until they found a comparison that provided the desired results. Mostly however, readers are not informed about failed comparisons.

Too often, studies have been published where the methods, results, and conclusions appear for the first time, allowing researchers to present the study as a neat package. Thus, researchers are able to change definitions and methods to fit data that have been collected. This "retro-fitting" in social sciences research is commonplace [30]. The Danish/American schizophrenia adoption studies provide an excellent example of this type of investigation. Both the definition of "schizophrenia" and the decision about which groups to compare were unduly influenced by the already-collected dataset [22]. Such post-hoc manipulation is however highly unethical [32]. The schizophrenia adoption study investigators [27] were therefore able to present their methods, results, and conclusions to the public for the first time in the same paper *as if by design*. By expanding the definition of "schizophrenia", they were able to conclude that their study provided evidence in favor of a genetic transmission of the condition. This conclusion was made despite having found zero cases of chronic schizophrenia amongst first-degree biological relatives of their group of index adoptees.

A priori registration with an ISSR could have prevented Kety and associates from coming to conclusions that did not flow from their dataset. Although registration would not become a legal requirement, studies not registered through ISSR would be viewed cautiously by other scientists. Information about methods submitted to ISSR would be kept confidential until publication. After publication, the information submitted to ISSR would become part of the public record.

As Medawar [33] observed, the format of traditional scientific papers is misleading (or "a fraud", as he put it), because "it misrepresents the process of thought that accompanied or gave rise to the work that is described in the paper" (p. 377). After writing his or her introduction and literature review, Medawar continued, the scientist writes a results section where "it is considered extremely bad form to discuss the significance of the results you're getting":

> You have to pretend that your mind is, so to speak, a virgin receptacle, an empty vessel, for information which floods into it from the external world for no reason for which you yourself have revealed. You reserve all appraisal of the scientific evidence until the "discussion" section, and in the discussion you adopt the ludicrous pretense of asking yourself if the information you've collected actually means anything; of asking yourself if any general truths are going to emerge from the contemplation of all the evidence you brandished in the section called "results" [33, p. 377].

Scientists are supposed to pretend that they have few preconceived ideas and passionate concerns, or a vested interest in the results of their study. This paper calls for the creation of ISSR, in part, because this is usually not the case.

(2) *When submitting papers reporting research experiments, the decision to publish should be made on the basis of a paper withholding the results and discussion sections.*

This idea has been proposed before [28,31,44]. In this framing, editorial acceptance of the paper would not be based on whether the results of the study were statistically significant or not. (This is particularly important for applied social science studies that require clinical or functional significance.) Rather, the "decision to accept" would be contingent on whether the experimental design was scientifically robust. Using this approach, the systematic bias against publishing studies that failed to reject the null hypothesis would be reduced, and replication studies, which are rare in the social sciences literature [29], would be encouraged. Similarly:

> There is a cardinal rule in experimental design that any decision regarding the treatment of data must be made *prior* to an inspection of the data. If this rule is extended to publication decisions, it follows that when an article is submitted to a journal for review, the data and the results should be withheld. This would insure that the decision to publish, or not to publish, would be unrelated to the outcome of research [44, p. 18].

Publication of a genuine scientific experiment should not depend on achievement of statistical significance, since both positive *and negative* results add to the body of scientific knowledge [16]. The bias in favor of statistically significant findings encourages researchers to look for such effects, so that their study can be accepted for publication [44]. Researchers may make distorted interpretations of their data in the context of research bias against negative findings.

(3) *Datasets should be made available to other researchers for independent analyses, particularly in studies that are difficult to replicate.*

Hypothetically, a team of social science investigators, through effort and ample funding, could study a small, unique subject group. The subject group could potentially help to the unravel the "nature/nurture" question. The subjects are so rare that, despite their importance, there were only a few similar previous studies ever recorded. Furthermore, subjects are unavailable for future study (that is, for replication study or follow-up). The investigators even acknowledge that prevailing social policies might preclude this type of work from ever being performed again. The researchers carefully guard their data from inspection, and refuse to release the data to other researchers for independent analysis. A few selected subjects (whose characteristics support the researchers' hypotheses) are presented to the press, although most outcomes remain unreported. The researchers publish several papers presenting only quantitative material on their subjects. These papers are later supplemented by several books from journalists that popularize the study. The audience is invited to radically reevaluate the importance of genetic influences on human psychological trait differences. Genetic explanations for individual differences are promoted.

This "hypothetical scenario" is precisely what has happened for the last 20 years in the Minnesota Study of Twins Reared Apart (MISTRA) of Bouchard and associates. The researchers have collected data on over 120 pairs of reared-apart twins, but have refused to allow independent investigators access to the data [8,24,45]. Scant information has been provided about the twins' histories. It is therefore impossible for other researchers to determine how *life experiences* (nurture) may have influenced the reported correlations (nature-type explanations). Yet from the distorted results and conclusions reported from the study, views have been misshapen about the supposed genetic contribution to human trait differences.

In a study of this type it should be mandatory to allow other investigators access to the raw data. There should be a consensus among social science investigators that conclusions would not be accepted from studies where investigators have withheld raw data from examination by other researchers. Bouchard and associates should have attended the warning from Arthur Jensen, who had written a preemptive attempt at damage control in the early days of the Burt scandal:

> Especially rare data, such as those of monozygotic twins reared apart, siblings from cousin matings, double first cousins, and the offspring of two mated pairs of monozygotic twins... should be published in full, along with complete descriptions of the tests or measurements and procedures. Perhaps this should be a general requirement for the publication of studies based on such valuable data, so that quantitative analytical techniques other than those used by the original author can be applied to the data by anyone who wishes [20, pp. 26–27].

It is understandable that some investigators have been reluctant to release the raw data from their published studies to third parties. Nonetheless, a system should be established within the original research protocol that minimizes the possibility of misuse by the third party.

(4) *The principal investigator's published conclusions will be followed by other alternative accounts and re-interpretations of the dataset.*

In the social sciences there is considerable emphasis on what researchers have *concluded* about their data. Many textbooks include multiple instances of unwarranted conclusions from limited datasets. Textbook writers often uncritically report research conclusions and may even report a study's conclusions from secondary sources [34]. In this way, urban myths and "folk devils" have been created from methods and results that never occurred, as exemplified by the descriptions of Watson's famous "Little Albert" case [38].

Unsurprisingly, most researchers conclude that their findings are consistent with their original hypothesis. Savan [39] noted the "pervasive… manner in which scientists can deliberately or, more often, unconsciously work in such a way that their conclusions are bound to support a particular position, policy, or action" (p. 26). Undoubtedly, there is a high degree of correlation between what researchers believe in advance that their studies will show, and what they concluded after the data were collected and analyzed. The investigators' conclusions, however, may not always flow logically from their data. For example, a group of leading behavior geneticists [35] found no significant correlation between the personality test scores of birthparents and their 245 adopted-away biological offspring. Nonetheless, they concluded that there were genetic influences on personality differences! The results of virtually any study are open to multiple interpretations. Hence the conclusions reached by the original researchers about their data are not necessarily more accurate or valid than alternative explanations. Moreover, non-experts from other fields can provide *more* plausible interpretations of data than the frequently biased views of the originating researchers.

The originating researchers should publish conclusions based on their results. Nonetheless, the simultaneous publication of alternative conclusions should occur more often. For example, journal Editors can send the final draft version of papers to researchers who will provide alternative explanations of the results. These alternative commentaries would be published at the end of the original study. This process already occurs with peer commentaries published routinely in prestigious international journals such as *Behavioral and Brain Science*. The current proposal would extend this idea, with between four and seven "third party" commentators. Also, the alternative conclusions would be published with equal status.

As well as providing alternative conclusions for readers of the original studies, the current proposal would promote more accurate reporting in textbooks. With a set of alternative conclusions, textbook authors would be more likely to report them, and textbook readers would be better informed about controversial studies with multiple explanations.

2. Conclusions

The current paper extends previous calls for fundamental changes in the ways social sciences research is completed and reported. Reform is needed for a more balanced reporting of research both to professionals and the public. Bias cannot be completely removed from the research and publication process. Researchers are human and tend to conduct research that will produce findings consistent with their belief systems. Savan [39] has noted that "we ask scientists to perform a virtually impossible feat – to remain resolutely skeptical about the theories they care most passionately about. We shouldn't be surprised if few scientists can live up to this ideal of behaviour" (p. 51). Science cannot be separated from the society and political structures in which it operates. As Hubbard and Wald observed, policy decisions "reflect the society in which the scientists live and work" [19, p. 7]. The myth of "disinterested science" should be laid to rest. Good science can coexist with bias, if safeguards exist to prevent unfounded beliefs and questionable conclusions from being converted into "facts". Unfortunately, this has sometimes occurred in the social sciences.

Adoption of the proposed four methods would help reduce fraud and would counteract research bias. The current methods of peer review and replication studies would be enhanced with the inclusion of these four additional proposals. Such radical action is needed to help eliminate biased research, and to promote high quality social science investigation.

References

[1] E. Altman and P. Hernon, eds, Research Misconduct: Issues, Implications, and Strategies, Ablex Publishing Corporation, Greenwich, CT, 1997.

[2] D. Atkinson, M. Furlong and B. Wampold, Statistical significance, reviewer evaluations, and the scientific process: Is there a (statistically) significant relationship?, *Journal of Counseling Psychology* **29** (1982), 189–194.

[3] C. Begg, Publication bias, in: *The Handbook of Research Synthesis*, H. Cooper and L. Hedges, eds, Russell Sage Foundation, New York, 1994, pp. 399–409.

[4] C. Begg and J. Berlin, Publication bias: A problem in interpreting medical data, *Journal of the Royal Statistical Society (Series A)* **151** (1988), 419–463.

[5] R. Bell, *Impure Science: Fraud, Compromise, and Political Influence in Scientific Research*, John Wiley and Sons, New York, 1992.

[6] M. Bohman, Some genetic aspects of alcoholism and criminality, *Archives of General Psychiatry* **35** (1978), 269–276.

[7] M. Bohman, C. Cloninger, S. Sigvardsson and A. von Knorring, Predisposition to petty criminality in Swedish adoptees, *Archives of General Psychiatry* **39** (1982), 1233–1241.

[8] T. Bouchard, D. Lykken, M. McGue, N. Segal and A. Tellegen, Response [Letter to the editor], *Science* **252** (1991), 191–192.

[9] M. Boyle, *Schizophrenia: A Scientific Delusion*?, Routledge, New York, 1990.

[10] W. Broad and N. Wade, *Betrayers of the Truth*, Simon and Schuster, New York, 1982.

[11] H. Cooper, K. DeNeve and K. Charlton, Finding the missing science: The fate of studies submitted for review by a human subjects committee, *Psychological Methods* **2** (1997), 447–452.

[12] K. Dickersin, Research registers, in: *The Handbook of Research Synthesis*, H. Cooper and L. Hedges, eds, Russell Sage Foundation, New York, 1994, pp. 71–83.

[13] P. Easterbrook, Directory of registries of clinical trials, *Statistics in Medicine* **11** (1992), 345–359.

[14] R. Fletcher, *Science, Ideology, and the Media: The Cyril Burt Scandal*, Transaction Publishers, New Brunswick, NJ, 1991.

[15] S. Gould, *The Mismeasure of Man*, W.W. Norton & Company, New York, 1981.

[16] A. Greenwald, Consequences of prejudice against the null hypothesis, *Psychological Bulletin* **82** (1975), 1–20.

[17] R. Haier, D. Rosenthal and P. Wender, MMPI assessment of psychopathology in the adopted-away offspring of schizophrenics, *Archives of General Psychiatry* **35** (1978), 171–175.

[18] L. Hearnshaw, *Cyril Burt: Psychologist*, Cornell University Press, Ithaca, NY, 1979.

[19] R. Hubbard and E. Wald, *Exploding the Gene Myth*, Beacon Press, Boston, 1993.

[20] A. Jensen, Kinship correlations reported by Sir Cyril Burt, *Behavior Genetics* **4** (1974), 1–28.

[21] J. Joseph, The equal environment assumption of the classical twin method: A critical analysis, *Journal of Mind and Behavior* **19** (1998), 325–358.

[22] J. Joseph, The genetic theory of schizophrenia: A critical overview, *Ethical Human Sciences and Services* **1** (1999), 119–145.

[23] J. Joseph, Not in their genes: A critical view of the genetics of attention-deficit hyperactivity disorder, *Developmental Review* (in press).

[24] J. Joseph, Separated twins and the genetics of personality differences: A critique, *American Journal of Psychology* (in press).

[25] R. Joynson, *The Burt Affair*, RKP, London, 1989.

[26] L. Kamin, *The Science and Politics of I.Q.*, Lawrence Erlbaum Associates, Potomac, MD, 1974.

[27] S. Kety, D. Rosenthal, P. Wender and F. Schulsinger, The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics, in: *The Transmission of Schizophrenia*, D. Rosenthal and S. Kety, eds, Pergamon Press, New York, 1968, pp. 345–362.

[28] J. Kupfersmid, Improving what is published, *American Psychologist* **43** (1988), 635–642.

[29] J. Kupfersmid and D. Wonderly, *An Author's Guide to Publishing Better Articles in Better Journals in the Behavioral Sciences*, Clinical Psychology Publishing Company, Brandon, VT, 1994.

[30] M. LaFollette, *Stealing Into Print*, University of California Press, Berkeley, CA, 1992.

[31] M. Mahoney, Publication prejudices: An experimental study of confirmatory bias in the peer review system, *Cognitive Therapy and Research* **1** (1977), 161–175.

[32] D. Martin, *Doing Psychology Experiments*, 4th edn, Brooks Cole, Pacific Grove, CA, 1996.

[33] P. Medawar, Is the scientific paper a fraud?, *The Listener* **70** (1963), 377–378.

[34] D. Paul, *The Politics of Heredity: Essays on Eugenics, Biomedicine, and the Nature-Nurture Debate*, State University of New York Press, Albany, NY, 1998.

[35] R. Plomin, R. Corley, A. Caspi, D. Fulker and J. DeFries, Adoption results for self-reported personality: Evidence for nonadditive genetic effects?, *Journal of Personality and Social Psychology* **75** (1998), 211–218.

[36] D. Rosenthal, P. Wender, S. Kety, J. Welner and F. Schulsinger, The adopted-away offspring of schizophrenics, *American Journal of Psychiatry* **128** (1971), 307–311.

[37] R. Rosenthal, The "file drawer problem" and tolerance for null results, *Psychological Bulletin* **86** (1979), 638–641.

[38] F. Samelson, J.B. Watson's Little Albert, Cyril Burt's twins, and the need for a critical science, *American Psychologist* **35** (1980), 619–625.

[39] B. Savan, *Science Under Siege: The Myth of Objectivity in Scientific Research*, CBC Enterprises, Montreal, 1988.

[40] R. Simes, Confronting publication bias: A cohort design for meta-analysis, *Statistics in Medicine* **6** (1987), 11–29.

[41] D. Sohn, Publication bias and the evaluation of psychotherapy efficacy in the reviews of the research literature, *Clinical Psychology Review* **16** (1996), 147–156.

[42] T. Sterling, Publication decisions and their possible effects on inferences drawn from tests of significance – or vice versa, *Journal of the American Statistical Association* **54** (1959), 30–34.

[43] T. Sterling, W. Rosenbaum and J. Weinkam, Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa, *The American Statistician* **49** (1995), 108–112.

[44] G. Walster and T. Cleary, A proposal for a new editorial policy in the social sciences, *The American Statistician* **26** (1970), 16–19.

[45] L. Wright, *Twins: And What They Tell Us About Who We Are*, John Wiley & Sons, New York, 1997.