Science and Engineering Ethics (2001) **7**, *7-14*

Comment

Research Practices in Need of Examination and Improvement

Harold Hillman

Keywords: scientific fraud, parafraud, intellectual honesty

INTRODUCTION

Belief in the function of the scientific method as a tool for improving the quality of research implies that the more accurately, carefully and honestly, a hypothesis is generated, the experiments are done, and the results interpreted, the more likely one is to arrive at more reliable hypotheses. Therefore, it behoves research workers to give serious attention to all these aspects of a project, not just those parts which are carried out in a laboratory. In the past, too little attention has been paid to the former aspects, with the result that much poor quality work has been published. Such publications have not only decreased the accuracy of our understanding of the phenomena being studied, but have increased the 'noise' which obscures the signals research workers wish to detect.

Definitions

A *hypothesis* is a unifying idea about a certain phenomenon, which may be a guess, extrapolation, assumption, concept, or mistake whose only value lies in its ability to predict the findings of future experiments or observations more accurately than another hypothesis it seeks to replace. A *finding* is a group of data—not always accurate—derived by experiment or observation. As Popper pointed out, an untested, untestifiable or unfalsifiable, hypothesis does not contribute certainty to knowledge. A finding is only necessarily true for the precise conditions of the particular experiment or observation; both of the latter may be distorted by inaccuracies, errors and artefacts. The aim of the research worker is to minimise the latter possible distortions.

Address for correspondence: Dr. Harold Hillman, M.B., M.R.C.S., Ph.D., Unity Laboratory of Applied Neurobiology, 3 Merrow Dene, 76 Epsom Road, Guildford, Surrey GU1 2BX, United Kingdom

1. RESEARCH PRACTICES IN NEED OF IMPROVEMENT

Among such practices are:

(i) not distinguishing between hypotheses and findings. For example, in biology, ideas which their original authors put forward as hypotheses are subsequently regarded as findings: (a) the bilipid nature of the cell membrane, proposed by Davson and Danielli:[1,2] (b) the mosaic model of the cell membrane, proposed by Singer and Nicholson;[3] (c) the chemical hypothesis of transmission, proposed by Katz:[4] (d) the involvement of G-proteins in cell signalling by Gilman:[5] (e) the programming of cell death by Kerr, Wylie & Currie.[6,7]

The originators of these hypotheses and their schools have brought a great deal of evidence in support of them, but much of it is compatible with alternative hypotheses, and does not bear crucially in support or contradiction of the different hypotheses. One can give four examples of this practice: the Singer-Nicholson hypothesis has replaced the Davidson-Danielli hypothesis of membrane structure and chemistry, but *both* of them are compatible with each other. The data leading to the *chemical* hypothesis of transmission across a synapse, are compatible with an electrical hypothesis of the same phenomenon.[8] In research into the genesis of cancer, multiple sclerosis, schizophrenia, etc, it is very difficult to know for certain if any biochemical changes found in diseased organs are the cause or the consequence of the illness. 'Axonal flow' is the passage of chemicals and granules down axons.[9,10] So far there have been no critical experiments, clarifying whether it is a particular biophysical phenomenon, or whether it is simply the addition of diffusion, metabolism, and movement of the body.

(ii) failure or unwillingness to identify crucial assumptions inherent in the thinking, procedures and interpretations, of experiments or observations. For example, the procedure of subcellular fractionation, by which the chemistry of different parts of the cells is elucidated, implies 24 assumptions; pharmacokinetics, which is the study of the distribution of drugs within the body, carries 22 assumptions; electron microscopy, by which the fine structure of cells is examined, implies 12; measurements of concentration of tissue constituents are referred to by such parameters as fresh weight, dry weight, incubated weight, tissue volume, tissue water, protein and DNA—each of which makes 4-5 separate assumptions and gives different answers.[11] In the Appendix are listed the inescapable assumptions inherent in the use of pharmacokinetic and subcellular procedures.

The reactions of cytologists to identification of these assumptions have been: to assert that such 'criticism' is nihilist; (as if it were not necessary to identify assumptions and test their warrantability); to allege that there are no other procedures for answering these questions; to suggest that the assumptions are made only by those who identify them; to deny that some of these assumptions are inherent in the procedures used; or to ignore the whole question:

(iii) not asking fundamental questions, sometimes on the assumption that the pioneers, the senior scientists and the predecessors have asked and answered them substantively, and so it is not necessary to raise them again. This ignores the circumstances that many advances have been made by examining anomalies and contradictions in the current consensual views;

(iv) selection of data to interpret and publish also occurs. In the physical and most biological sciences, it is considered unacceptable to select data. A manuscript should contain either all the material or data summarizing all of it. However, two kinds of research seem to allow selection. Firstly, there is a widespread practice in many areas of academic research of some research workers and drug companies' not publishing experiments which do not support their hypotheses. Secondly, in various electron microscopic procedures, so much data is produced that the microscopist has to select the pictures he or she submits for publication. In histology, it is also usually not considered necessary to use statistics in such research:

(v) accepting views on the basis of respect for the persons who believe them, rather than examining the evidence for the particular view oneself. The authorities may be the pioneers, the professors, the teachers, the supervisors, Nobel Prize winners or the authors of recommended textbooks.[12-14] This is almost universal in science, academia, theology and politics. It would be acceptable if those who create the consensuses were prepared to enter into free and sincere dialogue with all interested parties, and sometimes to change their minds:

(vi) research workers in biology carry out insufficient 'control' experiments, that is to say, testing the effects of their experimental procedures on the systems they are studying. For example, subcellular fractionation has been used for 60 years, but comprehensive control experiments have never been published on the effects of the complex experimental procedures themselves on the enzyme activities and their distribution in the subcellular fractions. Tissues have been grown in culture for 80 years, but there is a considerable dearth of publications comparing the biochemical and staining properties of the cultured tissues with the same properties in the parent tissues. Such control experiments are needed to justify the relevance of biological experiments in culture to the state of the living intact animal. Therefore, they may not be able to distinguish differences clearly, on the one hand, between effects on the two populations caused or modified by the experimental procedures, and, on the other hand, those effects caused by diseases, variables or characteristics of the populations in their native environments, unaffected by experimental procedures;

(vii) often inappropriate calibrations are used for the instruments and measuring differences. For example, if one wishes to compare the DNA, proteins, lipids or salts of the tissues, animals, plants, extracts, etc., one uses pure solutions of these substances for the calibrations, on the assumption that the extraction of the substances, the presence of other substances in their native tissues and the use of unnatural solvents for

the substances, would not affect the reading of the instruments. One should always do 'recovery' calibrations in which the substances calibrated have been added to the tissues, have passed through the whole experimental procedure, and are examined in the same or similar environments as they are found in the intact animal:

(viii) there is such a complicated range of difficulties in statistics, that it might be most useful to list some of the ground rules:[15-18]

(a) one cannot conclude from several series of similar experiments by different authors, each of which does not show a significant difference between two populations, that they altogether add up to a significant difference;

(b) different statistical tests examining the same data cannot produce significantly different degrees of significance;

(c) if one compares a hundred independent characteristics of two populations, 5% of them will be different by chance, with a probability of 0.05. Thus, if one goes on measuring many different characteristics of a population, or if one does not use all one's data in calculations, sooner or later, one will come across a run of results which will be apparently significantly different from the rest of the population. This may not be a truly biological difference, and can be tested by studying larger populations;

(d) many tests of significance of differences between two populations are based on the assumption that the variable measured shows a normal distribution in both populations. Sometimes the populations are too small to permit one to know whether or not the characteristic is normally distributed. If it is not, that particular statistical test may well be invalid;

(e) many statistical tests compare random populations. Of course, volunteers, observer-biased observations, and populations in which some values have been rejected on arbitrary grounds, are not.

2. FRAUD

Under this heading, one must consider: adding or subtracting values or numbers of the observations to those actually made; denying authorship to all who have shared the intellectual responsibility for the research; deliberately misquoting other authors: not reading submitted manuscripts or applications for funds sufficiently to assess their contribution to knowledge. By its nature, it is difficult to find out the extent of fraud, but it has been clearly documented to have been carried out by such famous scientists as Ptolemy, Mendel, Millikan, Burt, and many others more recently.[12,13,19-22]

3. PARAFRAUD

Parafraud constitutes another group of practices, which are frequently sins of omission, rather than commission. These include plagiarism, sometimes unwitting, of the ideas or experiments of other research workers including junior workers in one's own team; not preserving or being willing to show raw data to interested parties, including intellectual antagonists, so that calculations, measurements and calibrations can be checked, or

fraud detected; unwillingness to discuss details of published experiments; omitting from some publications such details that the experiments can be duplicated; ignoring findings which pre-date or contradict the author's own work; refusing to accept full intellectual responsibility for the publications of which one is a co-author; unwillingness to engage in dialogue with intellectual opponents, informally, at lectures or at meetings of learned societies; failure to answer proper questions in correspondence or at meetings; partisan refereeing of manuscripts for publication or applications for grants; preventing others with different viewpoints from putting them to learned society meetings, international conferences, or to the mass media; personal rudeness towards those with whom one disagrees.[14,22]

Unfortunately, like fraud, it is impossible to assess how widespread these practices are. There is no body monitoring them, and it is doubtful if the extent could be detected. Indeed, many of them are widely regarded by modern research workers as normal, if not admitted, weapons in the battle to advance their careers.

4. HOW RESEARCH CAN BE IMPROVED

Nevertheless, there are some measures which could be taken to improve research practices, prevent the distortions of truth and eliminate bad practices.
These include:

(i) teaching logic, semantics, the theory of knowledge, intellectual honesty, and statistics in all university courses;

(ii) encouraging students to challenge the accepted paradigms and to ask 'awkward' and fundamental questions;

(iii) deliberately encouraging a culture of academic integrity, including intellectual 'whistle blowing':

(iv) encouraging debates at learned societies on subjects in which there are several strongly felt different viewpoints:

(v) supporting research groups that have different views and approaches to the same problems;

(vi) insisting that all academics, however exalted, have a duty to enter into dialogue—including correspondence—about any research, experiments or observations, on topics about which they have published, spoken in public, refereed, or taught;

(vii) allowing those who submit manuscripts for publication or applications for grants to appear before open committees which consider them, both to answer and ask questions;

(viii) creating independent ombudsmen to consider whether authors who submit manuscripts for publication and applications for grants have been treated fairly by the committee considering them;

(ix) referees signing their recommendations to editors and grant giving bodies, and being prepared to justify them;

(x) requiring research workers or laboratories in which they did their experiments to preserve raw data so that other interested parties may have access to them in the future;

(xi) encouraging the dialogue between those who have different views on web sites;

(xii) encouraging the publication of papers in which authors admit that they have changed their minds or made mistakes, and the reasons for such admissions;

(xiii) inhibiting the use of personal abuse, appeals to authority, special pleading and the suppression of unpopular evidence in journals, refereeing, popular science newspapers, and at learned society and international meetings.

Poor quality research, fraud and parafraud, all have the effect of impeding progress towards truth, however it is defined, or however ephemeral it may turn out to be. Their origins are probably: those at the top of academia do not wish their power to be threatened by casting doubt on their own viewpoints; many research workers are poorly trained; junior research workers cannot risk their careers by questioning current consensuses; many research workers do not have the access to the literature or are ignorant of it. On the other hand, we live in an age of too much information, in which the noise is sometimes louder than the signals.

APPENDIX

Assumptions inherent in the use of two well known biochemical procedures.[11]

Pharmacokinetics

A drug is administered intraperitoneally to an animal to study its uptake into the different organs, its excretion and its metabolites. The following assumptions may be identified, that:

(i) the drug is always injected in the same site, whereas it may be subcutaneous intramuscular, intraperitoneal, into the alimentary tract, or into several of these compartments. The absorption from the different locations into the blood varies with the real site of injection;

(ii) the biochemical changes consequent on the stress to the animal do not alter the distribution in the circulation and the tissues of the drug;

(iii) the changes due to stress do not alter the affinity of the drug for the different tissues;

(iv) the changes due to stress do not alter the excretion by the kidney, liver, lungs or skin;

(v) the injection does not alter the appetite or thirst of the animal and, therefore, its nutritional status;

(vi) the injection does not alter the animal's behaviour in other ways such as its mobility, its sleeping habit, its resistance to infection or its temperature regulation, all of which could affect drug metabolism;

(vii) repeated injections do not produce accumulation or variation in any of the above effects;

(viii) the injection needle does not wound the animal;
[It will be rapidly appreciated that the injection of saline or the vehicle of the drug into control animals makes only a limited impact on these assumptions (v-viii)];

(ix) the removal of some animals of a batch does not produce stress to the other ones left in the cage;

(x) the circadian rhythms do not govern any of the previously mentioned effects;

(xi) agonal changes in blood biochemistry do not alter the distribution or metabolites of the drug;

(xii) the measurement of the drug in the tissues is not affected by the presence of the tissue, or that the extraction from each of the tissues is equal. This could only be tested by the use of a recovery method for calibration;

(xiii) the homogenisation does not affect the affinity of the binding of the drug by the tissue;

(xiv) other reagents besides the extractants do not affect the binding of the drug by the tissue;

(xv) other procedures such as centrifugation, filtration, washing etc. used in measurement of the drugs do not affect the binding of the drug by the tissue;
(xvi) the post-mortem changes do not affect the binding or any of the measurements;
(xvii) the particular metabolites or their proportions are not affected by the biochemical effects of stress;
(xviii) they are not affected by agonal changes;
(xix) the effect of the metabolites does not change after death;
(xx) the extraction of the drug from each of the homogenates or body fluids is equal or complete;
(xxi) the drug does not itself interfere with the measurement;
(xxii) the calibration of the drug is not different in the presence of tissue as in free solution. The use of standards passed all the way through the whole system warrants this assumption, but not the frequent use of standards used only with the final steps of the measurements.

Subcellular fractionation

A second example of assumptions inherent in the particular technique comes from measurement of the subcellular distribution of an enzyme.[23] The main inherent assumptions are that:

(i) the stress of handling and killing has no effect on the result of the experiment;
(ii) the agonal changes have no significant effect;
(iii) post-mortem changes have no significant effect;
(iv) cooling to room temperature, 0^{o}, -25^{o} or -196^{o}C has no significant irreversible effect;
(v) the enzyme activity of a homogenate decreases linearly with dilution;
(vi) the medium in which the tissue is homogenised, containing, for example, sucrose, edta, detergent or bile salts, does not alter the chemical activity significantly and irreversibly;
(vii) the enzyme activity measured finally is not changed significantly by the incomplete replacement of soluble constituents of the tissue which are lost on gross dilution, homogenisation and centrifugation in a quite different chemical environment;
(viii) movement during preparation of known co-factors, such as cations, and co-enzymes, or unknown ones, will not alter substantially the apparent localisation of enzyme activity as measured;
(ix) soluble materials originating from any compartment *in vivo* will not diffuse into the supernatant or become bound to another fraction during preparation, and thus be supposed to have originated in the location where it is found;
(x) no step in the preparation or lytic enzymes will render substances which were slightly soluble *in vivo* more soluble, and thus more diffusible, which may change their affinity for the different fractions;
(xi) the heat necessarily generated during homogenisation is so rapidly conducted away that the temperature does not rise sufficiently high to change enzyme activities irreversibly;
(xii) refrigerating the centrifuge diminishes temperature rise at the surface of particles being homogenised;
(xiii) enzyme activities are not irreversibly changed by pressure during homogenisation and centrifugation;
(xiv) the same amount of work is done on each different part of a centrifuge tube;
(xv) the same amount of heat is generated in different layers of the homogenate in media which have different viscosities;
(xvi) the extraction from each of the final fractions is equal and complete;
(xvii) a recovery of enzyme activity in all the subcellular fractions, added together of 60% to 130% of that in the initial crude homogenate, implies that the enzyme has not relocated;
(xviii) the enzyme preparation has no significant non-enzymic activity on the added substrate;
(xix) the enzyme activity in the unphysiological substrate mixture is approximately similar to that in the chemical environment *in vivo*;

(xx) the similarity in appearance on electron microscopy of, for example, a mitochondrion in a section to that in a mitochondrial fraction is evidence that its biochemical properties have not changed during fractionation;
(xxi) the microsomal fraction consists mainly of cell membranes and endoplasmic reticulum;
(xxii) the apparently high but unquantifiable incidence of a particular identifiable organelle in a particular fraction as seen by electron microscopy is evidence that the biochemical properties of that fraction are dominated by that organelle;
(xxiii) when the enzyme activity is referred to the protein content of a subcellular fraction, the protein can be measured satisfactorily;
(xxiv) the calibration method for protein in one subcellular fraction is applicable to other fractions; unless this has been tested for specifically.

REFERENCES

1. Gorter, M.D. & Grendel. F. (1925) On biomolecular layers of lipids on the chromocytes of blood. *Journal of Experimental Medicine* **41**: 439-443.
2. Davson, H. & Danielli, J.F. (1936) *The Permeability of Natural Membrane*, Blackwell, Oxford.
3. Singer, S.J. & Nicholson, G. (1972) The fluid mosaic model of the structure of the cell membrane, *Science* **175**: 720-731.
4. Katz. B. (1969) *The Release of Neural Transmitters*. Liverpool University Press, Liverpool.
5. Gilman, A.G. (1987) G-proteins: transducers of receptor-generated signals. *Annual Review of Biochemistry* **56**: 615-649.
6. Kerr, J.F.K., Wylie, A.H. & Currie, A.R. (1972) Apoptosis: A basic biological phenomenon with wide range implications in tissue kinetics. *British Journal of Cancer* **26**: 239-257.
7. Locksin, R.A. & Beaulaton, J. (1974) Programmed cell death. Life. *Science* **15**: 1549-1565.
8. Hillman, H. (1991) A new hypothesis for electrical transmission in the mammalian central nervous system. *Medical Hypotheses* **34**: 220-224.
9. Weiss, P. & Hiscoe, H.B. (1948) Experiments on the mechanism of nerve growth. *J. exp. Zool*. **107**: 315-395.
10. Weiss, D.G. (ed) (1982) *Axonal Transport*. Springer, Berlin.
11. Hillman, H. (1985) Some fundamental theoretical and practical problems associated with neurochemical techniques in mammalian studies. *Neurochemistry International, Supplement* **1**: 1-23.
12. Kohn, A. (1992) *False Prophets: Fraud and Error in Science and Medicine*. Blackwell, Oxford.
13. Broad, W. & Wade. N. (1993) *Betrayers of Truth: Fraud and Deceit in Science.* Simon Schuster, New York.
14. Moran, G. (1998.) *Silencing Scientists And Scholars in Other Fields*. Ablex Publishing Corporation. Greenwich, Connecticut.
15. Campbell, M.J. & Machin. D. (1993) *Medical Statistics, A Commonsense Approach*. Second edition. John Wiley, Chichester.
16. Armitage P. & Berry. G. (1994) *Statistical Methods in Medical Research*. Third edition. Blackwell, Oxford.
17. Rees, D.G. (1995) *Essential Statistics*. Third edition. Chapman & Hall, London.
18. Miller, H. (1995) *Descriptive Statistics*. Open Learning Foundation. Churchill Livingstone, Edinburgh.
19. Mahoney, M.J. (1976) *Scientist as Subject; The Psychological Imperative*. Ballinger, Cambridge (Mass).
20. Lock, S. & Wells. F. (1993) *Fraud and Misconduct in Medical Research*. British Medical Journal Publishing, London.
21. Grayson, L. (1995) *Scientific Deception*. British Library, London.
22. Lang, S. (1998) *Challenges*. Springer, New York.
23. Hillman, H. (1972) *Certainty and Uncertainty in Biochemical Techniques*. Surrey University Press, Henley on Thames, UK.