



8-1-1983

The Ombudsman: Cheating in Management Science

J. Scott Armstrong

University of Pennsylvania, armstrong@wharton.upenn.edu

Postprint version. Published in *Interfaces*, Volume 13, Issue 4, August 1983, pages 20-29.
Publisher URL: <http://www.aaai.org/AITopics/html/interfaces.html>

This paper is posted at Scholarly Commons. http://repository.upenn.edu/marketing_papers/112
For more information, please contact repository@pobox.upenn.edu.

The Ombudsman: Cheating in Management Science

Abstract

Honesty is vital to scientific work and, clearly, most scientists are honest. However, recent publicity about cases involving cheating, including cases of falsification of data and plagiarism, raises some questions: Is cheating a problem? Does it affect management science? Should anything be done?

Comments

Postprint version. Published in *Interfaces*, Volume 13, Issue 4, August 1983, pages 20-29.

Publisher URL: <http://www.aaii.org/AITopics/html/interfaces.html>

Published in *Interfaces*, 13 (4), 1983, 20-29.

Cheating in Management Science
(with Comments by M. K. Starr and M. J. Mahoney)

J. Scott Armstrong

The Wharton School, University of Pennsylvania

Honesty is vital to scientific work and, clearly, most scientists are honest. However, recent publicity about cases involving cheating, including cases of falsification of data and plagiarism, raises some questions: Is cheating a problem? Does it affect management science? Should anything be done?

My interest in scientific cheating began years ago, when I was responsible for the marketing research on a government-sponsored research contract. The purpose of the project was to examine an approach to urban transportation, which I will call the Urban Transit Plan (UTP). The government supported this research because of the potential social benefits if commuters were to switch from their automobiles to the UTP. My initial studies showed that more commuters would switch to the UTP from the existing mass transit than would switch from automobiles. Thus, instead of producing social benefits, the UTP would increase social costs. So what happened? First, I was asked by the project leader to "check" my results. What I inferred from this request was that I should not report these results, but should do studies until I obtained the desired results. I commissioned additional independent studies, but these supported the initial findings. Subsequently, I learned from informal sources that my reports were "lost" when the project leader reported to the government agency. Despite my objections to the project leader, additional reports were also lost. Therefore, I took action to ensure that the government received the reports. My action led to "peer review" whereby the 15 members of the project team, mostly university faculty, were asked by the project leader whether I should be fired. Only two faculty members voted that I should not be fired; however, they took such a strong stand in my favor that I was allowed to remain.

I have felt pressure to cheat on other projects in management science. Furthermore, I have heard stories of colleagues who experienced similar pressures. I have observed what seemed to be cases of cheating, and others have told me of cases that appeared to involve dishonesty of one sort or another. None of these cases were publicly revealed. My own efforts to report cheating on the UTP project apparently were of little interest to the government sponsor.

Recently, the number of publications on cheating in science has grown at a dramatic rate. A summary of some cases of scientific cheating is provided in Broad and Wade's *Betrayers of the Truth* [1982]. This provides vivid descriptions of 34 cases of alleged fraud in science. Almost all of these cases were identified since 1950, and over half have arisen in the last five years. Some of these cases refer to earlier work, dating as far back as Hipparchus in the second century BC. They include such famous scientists as Ptolemy, Galileo, Newton, Bernoulli, Dalton, Mendel, and Burt.

An analysis of the *Reader's Guide* from 1900 to the present supports the picture in Broad and Wade. We found little discussion of cheating prior to 1950. Some articles appeared between 1950 and 1974, and the number of papers has increased dramatically since 1975.

Does this mean that cheating is a serious problem? Initially I thought it did. After reviewing the literature and discussing the problem with others, I have concluded that I was partially wrong. Science is not threatened by cheating; but *scientists* are.

My initial conclusion was wrong because I fell prey to two common errors, even though I was previously aware of these errors. These errors are known in the research literature as "availability" and "vivid examples."

"Availability" states that events that are easy to recall are perceived as more likely. (For recent studies on availability, see Kahneman, Slovic, and Tversky [1982], part IV.) Projects that involved cheating are easier to

remember than projects that did not involve such conduct. Therefore, I perceived cheating as common. However, when I listed the projects that I had worked on over the last 23 years, the proportion that involved suspected cheating was smaller than I expected (less than five percent).

The effect of “vivid examples” was demonstrated by Hamill, Wilson and Nisbett [1980]. Subjects presented with vivid information about a single event generalized from this to the total population. In contrast, relevant statistical information on the population had little impact on the subjects’ estimates. Surprisingly, the vivid example had the same impact even when the subjects were told that it was atypical! My own experience combined with the vivid examples in Broad and Wade deceived me initially. This error was pointed out by my peers when they read early drafts of this paper. Consider the evidence in Broad and Wade, they said. Certainly no one would think that science is completely free of cheating. We now have evidence that cheating is likely to have occurred in 34 cases. That should not substantially affect your initial estimates of the frequency of cheating given the many thousands of scientists who have been working over the centuries.

Estimates of cheating could be obtained by systematic audits of samples of projects, but we have little information of this type. One exception is the Food and Drug Administration claim that 16 of the 50 physicians audited between 1967 and 1973 submitted false data on drugs to the sponsoring companies and the government [Wade 1973]. This, however, pertained to such a narrow sample that, once again, we can’t generalize.

Is management science different? None of the cases in Broad and Wade or in our review of the *Reader’s Guide* involved management science. In fact, I could find only one publicized case of alleged cheating in management science.

To find out how others perceive the problem of cheating in the management sciences, I sent a short questionnaire to the editors of 15 management science journals in December 1982. The journals were: *Academy of Management Journal*, *Academy of Management Review*, *Decision Sciences*, *Interfaces*, *Journal of Accounting Research*, *Journal of Business*, *Journal of Business Research*, *Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, *Journal of the Operational Research Society*, *Management Science*, *Marketing Science*, *Omega*, and *Operations Research*. I received 13 replies after a single mailing. In general, these editors were not aware of cheating in papers submitted to their journals. Four respondents were concerned about plagiarism, but estimated its frequency as low. (One editor, however, remarked that “ideoplagiarism,” stealing from oneself, is common.) Ten of the 13 respondents did not expect the problem of cheating to increase in the future.

Why Some Scientists Cheat

By examining the system of extrinsic rewards and punishments in science we can understand the pressures to cheat.

Casual observation indicates that scientific rewards often seem to go to those whose research reflects their advocacy. Mitroff [1972] presented evidence suggesting that many of the best known scientists often departed from objectivity in their attempts to support their theories.

Advocacy makes cheating attractive. When the facts do not conform to the favored theory, an advocate is motivated to change the facts. Many of the cases cited in Broad and Wade [1982] were of this type. Scientists try also respond to pressure from a project leader. For example, Smith [1982] presented evidence that the government persuaded some scientists involved with a 1953 Atomic Energy Commission study to suppress evidence about the effect fallout had on sheep. Roth [1966] suggested that research assistants may try to “help” their project leader by producing desirable results, even where the project leader exerted no overt pressure.

Because few cases of cheating have been reported, scientists may believe that cheating is unlikely to be detected. Or, when cheating is alleged, it may seem unlikely to be proven. The burden of proof is upon the accuser and judging from the non-probability survey by St. James-Roberts [1976], proving that cheating has occurred may be difficult. But of 185 cases of alleged cheating, the suspect was dismissed in only 10%, and many suspects were later promoted. But we do not know the percentage of those cases in which the charges were justified.

Changes to Reduce Cheating

Are major changes needed to protect science? I don't think so. Cheating has damaged science very little. Current safeguards, in particular replication, provide protection from serious scientific errors. This is the belief held by many of the senior scientists I consulted in preparing this paper, and my attempts to play devil's advocate produced little contrary evidence.

Although cheating does not seem to harm science, the published cases indicate that it does harm scientists. The cost to the whistleblower, as well as to the person who is charged, was immense in many cases. For example, Kammerer announced his impending suicide in his letter admitting to fraud [Zirkle 1954].

Instead of major changes, I suggest a few minor modifications to strengthen the existing safeguards. Furthermore, I think that attention should be given, not to making cheating more costly, but to making sure that the rewards do not encourage it. I have suggestions for each of the key participants in the process: the potential cheaters, the potential whistleblowers, the institutions, and the journals.

Potential Cheaters

Read Broad and Wade [1982] and other accounts of alleged cheating to increase your perception of the risk involved. Remember also that an informal network exists which evaluates stories about alleged cases of cheating in management science.

Potential Whistleblowers

The form of cheating most easily dealt with is plagiarism. If you are positive that someone has stolen your work, make your claim to the work in a way that treats the theft as a mistake. Forget about intent. If your damages are large, and you can prove them, you may have to see a lawyer. In short, worry about your reward and ignore punishment.

If you spot what you think is fraudulent data, my advice is: *Do nothing about the cheating*, because ...

1. your perceptions may be incorrect and you may falsely damage someone's reputation (How could I prove that my *perceptions* in the UTP case were realistic?),
2. the burden of proof will be on you. Because you have no rights of discovery from the accused, proof will be difficult (How could I prove that the reports in the UTP study were lost on purpose?), and
3. the penalties inflicted on you as a whistleblower may be severe – even, it appears, if you are correct. For serious cases, accept the possibility that you will lose your job (I almost lost my job over the UTP case without even making a public accusation.)

The published cases demonstrate these observations.

The most important reason for caution, however, is that the normal safeguards of science are sufficient. Whistleblowing is usually not needed. Surprising results about important topics will be questioned and attempts will be made to replicate them.

Instead of accusing someone of cheating, the whistleblower can follow accepted practice and try to replicate the study by following procedures outlined in the original study, and then trying to publish the results.

There are exceptions to all generalizations. For example, if whistleblowing will prevent serious harm to society, then you should act. Some cases of consumer product research (for example, in the pharmaceutical industry) fit

this description. Society has gained a lot from the heroic efforts of whistleblowers, as demonstrated in the cases described by Nader, Petkas, and Blackwell [1972]. Even in such cases, whistleblowers must be willing to sacrifice their jobs to protect society.

Institutions

Institutions should have policies for dealing with scientific cheating even though cheating is rare. According to Broad [1982], few institutions have such policies. However, in the wake of notorious cheating cases, Harvard Medical School and Yale University have taken steps to develop policies [*Harvard Gazette* 1982; Yale University 1982].

Yale believes that it is important to avoid regulation. They offer a gentle approach to whistleblowing, one that allows the cases to be handled with little publicity, and with little risk to the whistleblower:

All scholars have an obligation to disclose what they believe in good faith to be well-founded suspicions of academic fraud. Allegations of fraud must, of course, be made with great caution; yet those who come forward with such allegations must understand that the University respects the honest exercise of their judgment. At the same time, the rights of those whose work product is questioned must also be scrupulously protected, all in accord with a process that responds to such allegations with the utmost care, diligence, sensitivity, and respect for the rights of all concerned.

When institutions, rather than individuals, have handled cases of cheating, the whistleblowers received better treatment.

Such policies can help deal with cheating cases without the damage that public disclosure can inflict, even when charges are unsupported. (In retrospect, I think the government agency in my UTP case, by keeping the matter private, acted in everyone's best interests. Furthermore, they quickly cancelled their funding for the project on the grounds that the social benefits had not been demonstrated.)

Journals

To avoid the pressures of advocacy, journals could give preference to publishing papers that follow the method of multiple hypotheses. As suggested by Chamberlin [1890], this practice of examining all relevant hypotheses changes the role of the scientist. Instead of acting as an advocate for a particular hypothesis, the researcher obtains evidence to examine the relative merits of all the hypotheses. This should reduce the pressure to cheat on behalf of a favored hypothesis. Judging from my survey of papers published in *Management Science* [Armstrong 1979], however, only a small percentage of the papers used such a strategy.

Most editors feel responsible for ensuring that the papers they publish do not contain fraudulent data. DuShane, et al. [1961] apologized, as editors, because they had published a paper in *Science* that was later found to be based on fraudulent data.

According to my survey, most editors believe that the most effective way to handle cheating is to expose it by mentioning it in editorials, printing retractions, and notifying concerned people. I suggest that editors do this without making allegations of cheating. Most of us make mistakes in scientific work and it is difficult to be sure that others' mistakes are intentional.

Of course, in cases where fraud is admitted, or an error clearly demonstrated, a retraction should be published. The American Association for the Advancement of Science has taken the lead in this by publishing descriptions of such cases in their journal, *Science*.

The primary safeguard against cheating, as well as against mistakes or unusual events or false conclusions, is replication. Journals can strengthen replication by enforcing rules on full disclosure of methods and data and by increasing the acceptance rate for replication studies.

Authors are frequently frustrated in their attempts to obtain data on studies. Wolins [1962] was able to obtain information for replication from only 11 of 37 authors who had published in psychology journals. Wolin's study was extended some years later by Craig and Reese [1973]; they obtained supporting materials from about half of the authors in major psychology journals. Reid, et al. [1982] found that half of the 99 authors of papers in academic marketing journals that they contacted were unwilling or unable to make the requested materials available for replication.

Full disclosure, however, is expensive because of the cost of journal space and other costs such as duplicating and storage. Some editors report that for this and other reasons authors resist full disclosure. Nevertheless, some editors, like Albert Madansky of the *Journal of Business*, stressed disclosure. (I was impressed when, as a potential author for the *Journal of Business*, I was asked to submit the list of all the names that I had attempted to contact in a mail survey of experts.)

Formal policies on full disclosure might help. The *Journal of Forecasting*, drawing upon the experience of older journals, publishes the following guidelines in each issue:

For empirical studies the description of the method and the data should, ideally, be sufficient to allow for replication. Where this is not possible because of space restrictions, the authors should submit to the editors a complete description of the method and a copy of all data used. If this is not feasible because of cost, include summaries and samples of the data. Authors should also be prepared to make information available to interested readers. The *Journal of Forecasting* encourages replication studies.

This policy has been of demonstrable value, although it is not sufficient in all cases.

Replication is vital to science. If work cannot be replicated, it has little value. Attempts to replicate studies have frequently failed to confirm the original findings. Reid, Soley and Wimmer [1981], found that 40 percent of the 30 published replications they examined for advertising research conflicted with the original results. Wolin [1962] attempted to replicate seven studies in psychology and found gross miscalculation in three (again about 40 percent). Nevertheless, replication studies are not highly esteemed: In the Keir, Tolliver, and Petree [1977] survey of management and social science journals, 71 percent of the 301 editors and advisory board members of journals reported a reduced likelihood of accepting for publication a competent paper that was "a direct replication of an original study recently published in your journal (that) adds no new dimension to theory." Of course, some observers suggest that few studies are worth replicating; therefore, one cannot expect many replications to be published.

It is not surprising, then, that few replication attempts in the social sciences are published. Bozarth and Roberts [1972], in their survey of 1,334 papers published in psychology journals, found that less than one percent were replications of previously published papers. Reid, Soley, and Wimmer [1981] examined 501 papers in advertising research and concluded that less than one percent were direct replications and five percent were extensions.

A Modest Proposal for the Ombudsman Column

When I started work on this paper, I thought the Ombudsman column might be able to help prevent cheating. I considered many possible solutions, but for each one, colleagues were able to envision unintended consequences that made the costs outweigh the benefits. Eventually, I decided to forget about cheating and concentrate on replication. This eliminates worry about false perceptions, motives, libel suits, or unjustly damaged reputations.

If the results of a study cannot be replicated, there is no implication of cheating. It is likely to be the result of statistical variation or mistakes. Nor is it logical to assume that the rate of cheating is higher in a “failed to replicate” sample than in the general population of published papers. (This may not be intuitively obvious, so I will explain: Assume that 1 percent of scientists cheat, but that they always cheat to obtain results that are not surprising. Assume that 5 percent of the studies are surprising. Now assume that, being suspicious of new things, we take a sample of surprising studies truly. Of studies in this sample that fail to yield the same results, none would involve cheating.)

It seems valuable to encourage replication whether it has any impact on cheating or not. To increase information about such studies, I offer to print, in the Ombudsman, reports of replication attempts that have not yet been published. (I will also list cases where authors were unable to obtain sufficient information.) That an original study appears on the replication list can be viewed favorably by its authors, inasmuch as someone thought the study important enough to try to replicate it.

Listing will be made only at the request of authors. They should send me the title of the replication, a brief summary, and a copy of the paper. No paper that contains an allegation of fraud will be listed.

Acknowledgements

Cathy Lasowski assisted in the literature search and provided useful comments. Regina Loro typed the six drafts that led to this paper.

References

Armstrong, J. Scott (1979), “Advocacy and objectivity in science,” *Management Science*, 25 (5), 423-428.

Bozarth, H. D. and R. R. Roberts, Jr. (1972), “Signifying significant significance,” *American Psychologist*, 27 (8), 774-775.

Broad, William J. (1982), “Yale announces plan to handle charges of fraud,” *Science*, 218, No. 4567 (October 1), 37.

Broad, William and Nicholas Wade (1982), *Betrayers of the Truth: Fraud and Deceit in the Halls of Science*. New York: Simon and Schuster.

Chamberlin, T. C. (1890), “The method of multiple working hypotheses,” reprinted in *Science*, 148, No. 3671 (May 7), 1965, 754-759.

Craig, James R. and Sandra C. Reese (1973), “Retention of raw data: A problem revisited,” *American Psychologist*, 28 (8), 723.

DuShane, G. et al. (1961), “An unfortunate event,” *Science*, 134, No. 3483, 945-946.

Harvard Gazette (1982), “Medical school dean releases report on falsification of research data,” January 29, pp. 1 and 11-12.

Hamill, Ruth, T. D. Wilson, R. E. Nisbett, (1980), “Insensitivity to sample bias: Generalizing from atypical cases,” *Journal of Personality and Social Psychology*, 39 (4), 578-589.

Kahneman, Daniel, P. Slovic, and A. Tversky (1982), *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

- Kerr, Steven, J. Tolliver, and D. Petree (1977), "Manuscript characteristics which influence acceptance for management and social science journals," *Academy of Management Journal*, 20 (1), 132-141.
- Mitroff, Ian (1972), "The myth of objectivity, or why science needs a new psychology of science," *Science*, 18 (3), B613-B618.
- Nader, Ralph, P. Petkas, and Kate Blackwell (1972), *Whistle Blowing: The Report of the Conference on Professional Responsibility*. New York: Grossman Publishing.
- Reid, Leonard N., L. C. Soley, and R. D. Wimmer (1981), "Replication in advertising research: 1977, 1978, 1979," *Journal of Advertising*, 10 (1), 3-13.
- Reid, Leonard N., H. J. Rotfield, and R. D. Wimmer (1982), "How researchers respond to replication requests," *Journal of Consumer Research*, 9 (2), 216-218.
- Samelson, Franz (1980), "J. B. Watson's Little Albert, Cyril Burt's Twins and the need for a critical science," *American Psychologist*, 35 (7), 619-625.
- South, R. Jeffrey (1982), "Scientists implicated in atom test deception," *Science*, 218, No. 4572 (Nov. 5), 545-547.
- St. James-Roberts, Ian (1976), "Cheating in science," *New Scientist*, 72, No. 1028 (November 25), 466-469.
- Wade, Nicholas (1973), "Physicians who falsify drug data," *Science*, 180, No. 4090 (June 8), 1038.
- Wolins, Leroy (1962), "Responsibility for raw data," *American Psychologist*, 17 (9), 657-658.
- Yale University, *Policy Statement on Collaborative Research*, May 23, 1982
- Zirkle, Conway (1954), "Citation of fraudulent data," *Science*, 120 (August 6), 189-190.

A Comment from Martin R. Starr, former editor-in-chief, *Management Science*

How many papers in MS lend themselves to replication? Would you replicate an estimate of computer time to run an algorithm? The above is intended to suggest that something more is needed to explain what you mean by replication. Your conclusions were unexpected based on the survey's bent. I applaud the difference.

A Comment from Michael J. Mahoney, The Pennsylvania State University

The topic, "Cheating in Management Science," is not a new one in the sociology and psychology of science, but it has received increasing attention in the last few years – perhaps, in part, because of the sensationalism surrounding some of the more blatant instances of data fabrication and misrepresentation.

The discussion of motivations for cheating, however, is one which I think deserves more extensive reflection. While I concur with professor Armstrong that the system of extrinsic rewards and punishments in science offers some telling information about why scientists might be tempted to cheat, I think that the matter goes well beyond the extrinsic contingencies. I believe that the vast majority of scientists conduct their research with integrity and a genuine appreciation for their role in the collective process of inquiry. The blatant cases of data fabrication, data suppression, and misrepresentation of methodology are – in my opinion – fortunately rare. Of greater concern are the less blatant (and probably less conscious) practices of "massaging the data," minor lapses of memory, and a wide assortment of "human factors" that probably have a far greater impact on the overall quality of our knowledge

than the sensationalized incidents. As has been rather extensively documented, the boundaries between “objective” and “subjective” processes in science are anything but sharp [Mahoney 1976, 1979]. The distinction between “honest” and “dishonest” research is similarly cloudy. Although conscious and intentional cheating is (hopefully) a very rare phenomenon in science, unconscious and unintentional distortions of methodology and data analyses are probably more common than we yet appreciate.

As to ways to reduce the likelihood of cheating *in* science, I could not be in stronger agreement about the importance of replication as an advocated policy in science. Depending upon whose estimate one uses, replication is relatively rare in the social sciences – probably accounting for less than 5% of all published articles [Mahoney, in press]. Since replicational studies are unlikely to be published – and since publication is central to the power politics and epistemics of modern science – it is unlikely that replicative research will increase unless we do something to encourage it. Its main function is not to “police our ranks” for occasional cheating, but rather to add further information on the parameters and reliability of a reported phenomenon. I commend professor Armstrong’s “modest proposal” for reporting attempts to replicate prior research.

As to his recommendation that one “do nothing” about suspected instances of cheating, I have mixed reactions. On the one hand, I think he is being realistic and practical in noting that “whistleblowers” are likely to get in trouble for their efforts. However, it is sad that we do not have mechanisms to openly investigate suspected violations of integrity.

Finally, I think this topic demands that we analyze the psychosocial processes operative in developing scientific knowledge. What is it about contemporary science that seems to encourage (rather than discourage) the “white lies” that some fear to be rampant? I think that our energies would be better invested in studying the factors which encourage cheating (rather than the interventions we might pursue to discourage and minimize cheating) [Mahoney 1976, 1979]. I believe that a substantial part of this problem can be attributed to the great weight placed on research publication in academic hiring, promotion, tenure, and awarding grants. Instead of serving as a method of communication for scientists, research publications dictate who and what will be funded and certified as “worthwhile” in contemporary science. My own tentative suggestions for reform are that we reappraise (1) the emphasis placed on quantity of publication in all personnel and funding decisions, and (2) our tenure system in academic science. These fundamental issues lie at the core of the social system of science; we need to develop a deeper appreciation of their effects on the overall quality of our work.

Although I concur with Professor Armstrong on the value of replication, I do not think replication will solve the problem of cheating in science. Conscious or unconscious, cheating is functional behavior. It is a means toward an end – and in academia, that end seems to be publication, personal advancement, and the attendant social reinforcements. Perhaps it is time that we “re-search” our basic assumptions about the processes that produce scientific progress and try to differentiate our concepts of “scientific contribution” and “publication.”

References

- Mahoney, M. J. (1976), *Scientists as Subject: The Psychological Imperative*. Cambridge, MA: Ballinger.
- Mahoney, M. J. (1979), “Psychology of the scientist: An evaluative review,” *Social Studies of Science*, 9 (3), 349-375.
- Mahoney, M. J. (in press), “Publication, power, and scientific progress: Notes on opening the knowledge market,” *Journal of Counseling Psychology*.