



May 1979

Advocacy and Objectivity in Science

J. Scott Armstrong

University of Pennsylvania, armstrong@wharton.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/marketing_papers

Recommended Citation

Armstrong, J. S. (1979). Advocacy and Objectivity in Science. Retrieved from http://repository.upenn.edu/marketing_papers/118

Postprint version. Published in *Management Science*, Volume 25, Issue 5, May 1979, pages 423-428.

Publisher URL: <http://mansci.pubs.informs.org/>

The author asserts his right to include this material in ScholarlyCommons@Penn.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/marketing_papers/118

For more information, please contact libraryrepository@pobox.upenn.edu.

Advocacy and Objectivity in Science

Abstract

Three strategies for scientific research in management are examined: advocacy, induction, and multiple hypotheses. Advocacy of a single dominant hypothesis is efficient, but biased. Induction is not biased, but it is inefficient. The multiple hypotheses strategy seems to be both efficient and unbiased. Despite its apparent lack of objectivity, most management scientists use advocacy. For example, 2/3 of the papers published in a sampling of issues of *Management Science* (1955-1976) used advocacy. A review of the published empirical evidence indicates that advocacy reduces the objectivity of the scientists. No evidence was found to suggest that this lack of objectivity could be overcome by a "marketplace for ideas" (i.e., publication for peer review). It is recommended that the method of multiple hypotheses be used.

Comments

Postprint version. Published in *Management Science*, Volume 25, Issue 5, May 1979, pages 423-428.

Publisher URL: <http://mansci.pubs.informs.org/>

The author asserts his right to include this material in ScholarlyCommons@Penn.

Advocacy and Objectivity in Science

J. Scott Armstrong

Abstract

Three strategies for scientific research in management are examined: advocacy, induction, and multiple hypotheses. Advocacy of a single dominant hypothesis is efficient, but biased. Induction is not biased, but it is inefficient. The multiple hypotheses strategy seems to be both efficient and unbiased. Despite its apparent lack of objectivity, most management scientists use advocacy. For example, 2/3 of the papers published in a sampling of issues of *Management Science* (1955-1976) used advocacy. A review of the published empirical evidence indicates that advocacy reduces the objectivity of the scientists. No evidence was found to suggest that this lack of objectivity could be overcome by a "marketplace for ideas" (i.e., publication for peer review). It is recommended that the method of multiple hypotheses be used.

1. Introduction

Objectivity is the foundation of scientific work. The objective scientist is looked upon as the ideal to strive for. At least, that is the popular conception of the ideal scientist.

Ian Mitroff [21] challenged this idealized conception. In a study of space scientists, Mitroff found that the scientists with the highest prestige did not live up to this ideal. Nor did they even try to live up to it. Instead, they adopted a dominant hypothesis and became advocates of this hypothesis. They sought evidence to confirm their hypothesis, and they presented this evidence in such a way as to put their hypothesis in the most favorable light. These findings suggested that successful scientists are not objective; they are biased advocates.

What makes advocacy work, says Mitroff, is the marketplace for ideas. Different scientists advocate their hypotheses and, over time, the most useful hypotheses survive.

Rather than viewing advocacy as a liability, Mitroff sees it as a legitimate and perhaps desirable approach to science. Mitroff's viewpoint struck a responsive chord among many scientists. On the other hand, some people, including myself, continue to view advocacy as a poor approach to science [1], [22].

In this paper, the advocacy strategy is compared with alternative approaches to science. Results are then presented on which scientific strategies are most popular in management science. Finally, published empirical evidence is examined to determine which scientific strategy should be used.

2. Advocacy and Alternatives to Advocacy

Does advocacy provide the best strategy for scientific research? To answer this question, it is necessary to develop a working definition for "best" and to outline some alternatives to advocacy.

"Best" is defined here by two dimensions: objectivity and efficiency. "Objectivity," the extent to which the researcher's methods are free of prejudice, has long been regarded as a characteristic of the scientific method and is the criterion of primary concern in this paper. "Efficiency" is concerned with achieving results with the least expenditure of time and money.

To decide on the most reasonable alternatives to advocacy, it is helpful to review three strategies for scientific research: induction (no hypothesis), advocacy (one dominant hypothesis), and multiple hypotheses (2 or more reasonable hypotheses).

The inductive strategy is based on careful and explicit observation. No prior hypotheses are formed. Data are collected and interpreted in an effort to develop useful hypotheses.

Science appeared to make a significant advance when advocacy was adopted. The single hypothesis provides a structure for the collection and analysis of the data. It allows for disconfirmation of the hypothesis; the data are collected to allow for a comparison between the dominant hypothesis and a simple null hypothesis.

Although advocacy improves efficiency by directing the scientist's efforts towards the testing of a single hypothesis, it also contributes to a loss of objectivity for the individual scientist [13]. The researcher often selects a simple, but unreasonable null hypothesis that offers no challenge to his favored hypothesis. For example, econometricians have published papers in which they defeat the null hypothesis that "the purchase of automobiles is unrelated to the income of potential buyers."

The most promising suggestion for retaining efficiency while avoiding the bias of the single dominant hypothesis is to select a second reasonable hypothesis. This could take the form of the null hypothesis except that it is selected for its reasonability instead of its simplicity (in the automobile example, a reasonable null hypothesis is that income elasticity is + 1.0).

This strategy of using at least two reasonable hypotheses was discussed in 1890 by Chamberlin [6]. He claimed that the use of the method of multiple hypotheses distinguishes sciences that progress rapidly from those that progress slowly. By adopting the method of multiple hypotheses, the role of the scientist is changed from advocate to experimenter. He is judged by his effectiveness in comparing the usefulness of various reasonable hypotheses. This method reduces bias because the scientist works with all of the leading hypotheses; they are all "his." Certainly he may have more confidence in one hypothesis than another; but he should obtain evidence to test each of these hypotheses. His job is to design critical tests for identifying the most useful hypothesis. He examines the competing hypotheses immediately, rather than waiting for an examination in the marketplace of ideas. Chamberlin's paper has been reprinted a number of times and has been interpreted by others (e.g. [23]).

Multiple hypotheses about the impact of the various scientific strategies are proposed in Table 1. This table ranks the objectivity and efficiency of the strategies. The method of multiple hypotheses, with both high objectivity and high efficiency, is expected to be superior to the other strategies.

Table 1
Hypothesized Rankings of Scientific Strategies
 (1 = most preferred ranking)

Strategy	Objectivity	Efficiency
Induction	1	3
Advocacy	3	1
Multiple Hypotheses	1	1

Does advocacy by different researchers promote objectivity? Although the publication of different viewpoints is desirable, Mitroff's argument [21, p. B615] that "Objectivity results from the heated, intense, and biased confrontation and struggle between the 'somewhat' biased ideas of somewhat biased individuals" is not intuitively obvious. Nor does it appear that different viewpoints receive a fair hearing. The use of referees by journals can be viewed as a type of censorship, and these censors are generally selected from those holding the currently dominant hypothesis. In Mahoney's [17] experiment, journal referees gave a high rating to relevance and methodology for a paper whose results were in agreement with the dominant hypothesis for their group. However, when the identical paper was used, but the results were reversed so as to be disconfirming, an equivalent sample of referees gave it a low rating on relevance and methodology. The referees were willing to accept the studies confirming the dominant hypothesis, but they rejected those that were disconfirming. This type of peer review does little to advance scientific knowledge. (Smart's study [26] is consistent with the Mahoney experiment: An examination of over 300 papers published in four psychological journals found that less than 10% reported negative results.)

Although scientists have been using the advocacy method, the efficiency of the process has not been self-evident, especially when one reviews the history of major advances in science (see [3]). Descriptions of outstanding advances in science indicate that the innovator is subjected to personal attacks, and his contribution is often ignored until after his detractors (and perhaps he himself) have died.

One proposal that might improve the objectivity and efficiency of the advocacy method is the use of a "science court" This court would ensure that different viewpoints are heard [4]. Furthermore, it would greatly reduce the time required for the presentation of alternative viewpoints. The procedure would be similar to that used in a court of law. Unfortunately, I could not find empirical evidence on the value of this strategy. (One version of a science court is illustrated in [2].)

3. The Popularity of Advocacy in Management Science

Those who believe in the "test of the marketplace" would ask what research strategies are used by management scientists. Presumably, scientists will tend to adopt the most useful methods in their work. At least, that is the argument based on the marketplace.

Mitroff [21] provided evidence that advocacy is a common strategy among scientists. This is consistent with historical reviews ([3] and [11]). Examples of advocacy from Lewis' *Arrowsmith* [16] and from Watson's *The Double Helix* [32] imply that advocacy has been around a long time and that scientists are indeed emotional in support of their hypotheses. Mahoney and Kimper [19], in a survey of scientists, found that the vast majority preferred experiments that confirmed their hypotheses.

To examine the situation in management science, a sample survey was made of papers that appeared in *Management Science*. Presumably, these published papers are representative of the best work done by management scientists. The survey covered 120 empirical papers appearing in selected years from 1955 to 1976. The papers were coded independently by four people to determine whether they used induction, advocacy, or multiple hypotheses. Agreement was found among at least three of the four coders for 80 of these papers. The other 40 papers were excluded due to unreliable coding.)¹

The results, summarized in the last column of Table 2, show that 64% of the studies used the advocacy strategy. Advocacy was substantially more popular than induction or multiple hypotheses.

Table 2
Classification of Papers from *Management Science*

Strategy	Percentage of Papers from:			Average Percentage
	1955-1959	1963-1966	1967-1979	
Induction	16	19	7	14
Advocacy	68	54	69	64
Multiple Hypotheses	16	27	24	22
Sample Size	(25)	(26)	(29)	(80)

¹ Three research assistants were each asked to select empirical papers (real or simulated data) from Management Science. One assistant started in 1976 and worked backwards in time taking all studies that had data until he located 40 empirical studies. A second assistant started in 1966 and also worked backwards to obtain 40 studies. The third assistant started in 1955 (the earliest volume in the series) and worked forward in time to obtain 40 studies. Each assistant coded each paper as to induction, advocacy, or multiple hypotheses. The papers were then blind-coded by the other two research assistants. For some papers there was no agreement, so they were dropped. On 48 papers, only two coders agreed on a rating so these were blind-coded by the author. Copies of the instructions to coders and a listing of the ratings can be obtained from the author.

Table 2 presents data over a span of about 20 years. The popularity (or lack of popularity) of the method of multiple hypotheses has remained relatively constant over this time: only 22% of the papers used multiple hypotheses. No significant differences nor trends were found. Advocacy continues to be the strategy favored by management scientists.

4. Empirical Evidence on Advocacy and Its Alternatives

A review of the literature revealed some evidence relevant to the hypotheses in Table 1. Most of the available research related to the issue of objectivity rather than efficiency.

Wason [28]-[31] examined objectivity. In his basic experiment, subjects were provided with a set of three numbers, 2-4-6, and were asked to identify the rule used to generate these numbers. Subjects could obtain information by generating their own sets of numbers and being told whether or not each new set conformed to the rule. The typical subject quickly selected a hypothesis; he then had difficulty in rejecting his hypothesis. He either did not want to or did not know how to disconfirm his hypothesis. Instead, he invested much energy to confirm his hypothesis. These studies suggest that the selection of a single dominant hypothesis will reduce objectivity.

DeMonbreun and Mahoney [10], in a further study with Wason's 2-4-6 problem, found that a subject usually maintained confidence in a hypothesis even when over 50% of the subject's predictions were wrong. The subjects tended to believe the feedback when it was positive, but disregarded it when it disconfirmed the hypothesis.

The 2-4-6 problem has also been used with a sample of scientists. Mahoney and DeMonbreun [18] found that psychologists and physical scientists selected experiments that confirmed their hypotheses. Less than 40% of the scientists tested their hypotheses by seeking disconfirming evidence; only 10% of their experiments sought disconfirming evidence.

In a study of how clinical psychologists use data to identify homosexuality, Chapman and Chapman [7] concluded that the single hypotheses leads to bias. Practicing psychologists were presented with contrived data and were asked to draw inferences on which variables were related to homosexuality. The data refuted much of what the psychologists believed; however, the psychologists ignored this disconfirming evidence and persisted in seeing the relationships they expected. That is, the psychologists did not revise their hypothesis when presented with disconfirming evidence.

The bias for a single dominant hypothesis shows up in other ways. Calculation errors made by a scientist tend to favor the scientist's hypothesis (see [14], [25]). The reactions of the subjects may be affected by the experimenter's bias (see [15], [25]). Sometimes the bias is so strong that it leads to cheating, as is the well-known cases of Isaac Newton in the physical sciences [33], Cyril Burt in psychology [27], and William Summerlin in biology [8], [9]. As stated in Maier's Law [20], "If the facts do not conform to the theory, they must be disposed of."

Many factors affect the bias introduced by scientists. Increases in the strength of the scientist's prior belief, the complexity of the situation, and the random error in the data all tend to increase the bias.

Actions by the scientist also affect bias. The mere act of selecting a single hypothesis will introduce bias-even when the scientist has no prior belief, the situation is not complex, and there is no error in the data. In Pruitt [24], subjects were presented with a sequence of lights and were asked to judge whether the sequence was mostly "red" or "green." If the subject selected a prior hypothesis (even though based on no information), this increased the amount of disconfirming evidence needed by the subject to change the hypothesis.

Geller and Pitz [12] conducted a study similar to Pruitt's except that poker chips were used. One bag contained 40% red chips and 60% white; the ratio was reversed in a second bag. If the subjects started with a single arbitrary hypothesis, disconfirming evidence did not initially cause them to reduce confidence; instead the initial reaction was an increase in their confidence in the face of this disconfirming evidence!

On the positive side, scientists can take actions to reduce bias. In a study similar to Pruitt's, Brody [5] provided subjects with paper and pencil and asked them to tabulate the results. In this artificial world composed of two hypotheses, this step helps to focus attention on the two reasonable hypotheses. The subjects, who were asked to select a prior hypothesis, did gain some initial confidence in the hypothesis, but they were able to adjust their beliefs as data became available. In other words, the added confidence in one hypothesis was not a biasing factor when data were tabulated to test multiple hypotheses.

5. Conclusions

Although the traditional ideal of the objective scientist has been challenged, little evidence exists to suggest that this ideal should be discarded in favor of advocacy. No evidence was found to support the statement that an objective approach to science can be achieved by having biased scientists present evidence on their favored hypotheses.

Unfortunately, our ideals conflict with our reward systems. We reward scientists who act as advocates. They are the ones who receive grants, who get promoted, who find it easiest to publish their findings, and who become the center of attention. Scientists are often identified by the theory that they advocate ("Professor Jones, the well known founder of the riddle theory"). Perhaps there is an analogy between selling theories and selling other items-like insurance and used cars. Successful salesmen use advocacy with these products also.

How can we make the practice of science conform more closely to the ideal of objectivity? Urging scientists to be more objective is of little value. However, we could promote the methods of multiple hypotheses:

1. Referees for journals could be asked to give preference to papers that use the method of multiple hypotheses.
2. Sponsors of research could ask for proposals that employ the method of multiple hypotheses.

Finally, it is important to note that the method of multiple hypotheses does not prevent scientists from reaching conclusions. They can even be enthusiastic about these conclusions.

Although advocacy is the strategy that is used and rewarded, the limited empirical evidence suggests that this is unfortunate. *The advocacy method should be abandoned, and multiple hypotheses should be used!*

References

1. Armstrong, J. Scott (1979), "Advocacy as a scientific strategy: The Mitroff myth," Working Paper, Dept. of Marketing, Wharton School, University of Pennsylvania, Philadelphia, Pa., 1979.
2. Armstrong, J. Scott (1978), "Econometric forecasting and the science court," *J. Business*, 51, 595-600.
3. Barber, Bernard (1961), "Resistance By Scientists To Scientific Discovery," *Science*, 134, 596-602.
4. Boffey, Philip M. (1976), "Experiment planned to test feasibility of a 'science court'," *Science*, 193, 129. See also the report by the Task Force of the Presidential Advisory Group on Anticipated Advances in Science and Technology in *Science*, 193, 653-656 and the letter by Earl Callen on 950-951 in the same volume of *Science*.
5. Brody, Nathan (1965), "The effect of commitment to correct and incorrect decisions on confidence in a sequential decision-task," *Amer. J. Psychology*, 78, 251-256.
6. Chamberlin, T. C. (1965), "The method of multiple working hypotheses," *Science*, 148, 754-759 (originally published in *Science* in 1890).

7. Chapman, Loren J. and Chapman, J. P. (1969), "Illusory correlation as an obstacle to the use of valid psychodiagnostic signs," *J. Abnormal Psychology*, 74, 271-280.
8. Culliton, Barbara J. (1974), "The Sloan-Kettering affair: A story without a hero," *Science* 184, 644-650.
9. Culliton, Barbara J. (1974), "The Sloan-Kettering affair (II): An uneasy resolution," *Science*, 184, 1154-1157.
10. Demonbreun, Bobby G. and Mahoney, Michael J. (1976), "The effect of data return patterns on confidence in an hypothesis," in Michael J. Mahoney, *Scientist as Subject*, Cambridge, MA: Ballinger, 181-186.
11. Feyerabend, Paul K. (1965), "Problems of Empiricism," in Robert G. Golodny (ed.), *Beyond the Edge of Certainty*. Englewood Cliffs, NJ: Prentice-Hall.
12. Geller, E. Scott and Pitz, Gordon F. (1968), "Confidence and Decision Speed In The Revision Of Opinion," *Organizational Behavior and Human Performance*, 3, 190-201.
13. Greenwald, Anthony G. (1975), "Consequences Of Prejudice Against The Null Hypothesis," *Psychological Bulletin*, 82, 1-20.
14. Johnson, Ronald W. and Adair, J. G. (1970), "The effects of systematic recording error vs. experimenter bias on latency of word association," *J. Experimental Res. in Personality*, 4, 270-275.
15. Kintz, B. L. et al. (1965), "The experimenter effect," *Psychological Bulletin*, 63, 223-232.
16. Lewis, Sinclair (1925), *Arrowsmith*. New York: Harcourt, Brace and World.
17. Mahoney, Michael J. (1977), "Publication prejudices: An experimental study of confirmatory bias in the peer review system," *Cognitive Therapy and Research*, 1, 161-175.
18. Mahoney, Michael J. and Demonbreun, B. G. (1977), "Psychology of the scientist: An analysis of problem-solving bias," *Cognitive Therapy and Research*, 1, 229-238.
19. Mahoney, Michael J. and Kimper, T. P. (1976), "From ethics to logic: A survey of scientists," in Michael J. Mahoney, *Scientist as Subject*. Cambridge, MA: Ballinger, 187-193.
20. Maier, N. R. F. (1960), "Maier's Law," *American Psychologist*, 15, 208-212.
21. Mitrofi, Ian (1972), "The myth of objectivity or why science needs a new psychology of science," *Management Science*, 18, B613-B618.
22. Operations Research (1971), "Guidelines for the practice of operations research," 19, 1123-1148.
23. Platt, John R. (1964), "Strong inference," *Science*, 146, 347-353.
24. Pruitt, D. G. (1961), "Informational requirements in making decisions," *American Journal of Psychology*, 74, 433-439.
25. Rosenthal, Robert and Rosnow, Ralph L. (1969), *Artifact in Behavioral Research*. New York: Academic Press.
26. Smart, Reginald (1964), "the importance of negative results in psychological research," *Canadian Psychologist*, 5a, 225-232.
27. Wade, Nicholas (1976), "IQ and heredity: Suspicion of fraud beclouds classic experiment," *Science*, 194, 916-919.

28. Wason, P. C. (1960), "On the failure to eliminate hypotheses to a conceptual task," *Quarterly Journal of Experimental Psychology*, 12, 129-140.
29. Wason, P. C. (1968), "On the failure to eliminate hypotheses – a second look," in P. C. Wason and P. N. Johnson-Laird (eds.), *Thinking and Reasoning*. Baltimore: Penguin, 165-174.
30. Wason, P. C. (1968), "Reasoning about a rule," *Quarterly Journal of Experimental Psychology*, 20, 273-281.
31. Wason, P. C. (1969), "Regression in reasoning?," *British Journal of Psychology*, 60, 471-480.
32. Watson, James D. (1968), *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. London: Weidenfeld.
33. Westfall, Richard (1973), "Newton and the fudge factor," *Science*, 179, 751-758.