



April 1980

Unintelligible Management Research and Academic Prestige

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. (1980). Unintelligible Management Research and Academic Prestige. Retrieved from http://repository.upenn.edu/marketing_papers/117

Postprint version. Published in *Interfaces*, Volume 10, Issue 2, April 1980, pages 80-86.

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

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/marketing_papers/117

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

Unintelligible Management Research and Academic Prestige

Abstract

Modest support was found for the "Dr. Fox Phenomenon": Management scientists gain prestige by unintelligible writing. A positive correlation (+0.7) was found between the prestige of 10 management journals and their "fog indices" (reading difficulty). Furthermore, 32 faculty members were asked to rate the prestige of four passages from management journals. The content of the passages was held constant while readability was varied. Those passages that were more difficult to read were rated higher in research competence.

Comments

Postprint version. Published in *Interfaces*, Volume 10, Issue 2, April 1980, pages 80-86.

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

Unintelligible Management Research and Academic Prestige

Reprinted with permission from *Interfaces*, Vol. 10, No. 2, April 1980, p. 80-86.

J. Scott Armstrong

The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

ABSTRACT. Modest support was found for the "Dr. Fox Phenomenon": Management scientists gain prestige by unintelligible writing. A positive correlation (+0.7) was found between the prestige of 10 management journals and their "fog indices" (reading difficulty). Furthermore, 32 faculty members were asked to rate the prestige of four passages from management journals. The content of the passages was held constant while readability was varied. Those passages that were more difficult to read were rated higher in research competence.

Introduction

Dr. Fox was an actor who looked distinguished and sounded authoritative. He was provided with a fictitious but impressive biography and was sent to lecture about a subject on which he knew nothing. The talk, "Mathematical Game Theory as Applied to Physician Education," was delivered on three occasions to a total of 55 people. One hour was allowed for the talk and 30 minutes for discussion. The audiences consisted of highly educated social workers, psychologists, psychiatrists, educators, and administrators. The lecture was comprised of double talk, meaningless words, false logic, contradictory statements, irrelevant humor, and meaningless references to unrelated topics. Judging from a questionnaire administered after the talk, the audience found Dr. Fox's lecture to be clear and stimulating. None of the subjects realized that the lecture was pure nonsense [Naftulin et al., 1973].

If an unintelligible communication is received from a legitimate source and if this communication claims to be in the recipient's area of expertise, recipients might assume that they are wasting their time because they receive no useful knowledge. In terms of knowledge, they

would be wasting their time. But their involvement in this activity may lead them to try to justify the time spent. Furthermore, the greater the unintelligibility, the greater the need to rationalize about the time spent (e.g., if you cannot understand a paper, it must be a high level paper). This might be called the *Dr. Fox hypothesis*: An unintelligible communication from a legitimate source in the recipient's area of expertise will increase the recipient's rating of the author's competence.

If the Dr. Fox hypothesis is valid, researchers who want to impress their colleagues should write less intelligible papers. Journals seeking respectability should publish less intelligible papers. Academic meetings should feature speakers who make little sense. This strategy would be beneficial for advancement by an individual researcher or by a journal. Its major drawback is that it does not promote the advancement of knowledge.

If one believes that academic communications should enhance knowledge, researchers should invest energy in developing understandable ways to present their findings. Academic conferences and journals should look for researchers who have interesting studies *and* can present them clearly. Other things being equal, researchers who are also good communicators should be rewarded more highly. This is called the "communication-for-knowledge hypothesis."

This paper contrasts the "communication-for-knowledge" and "Dr. Fox" hypotheses using written communications. Which hypothesis best describes the reaction of readers? Are they impressed by unintelligibility (Dr. Fox) or by clarity (communication-for-knowledge)?

Previous Research

The rational viewpoint favors the hypothesis that the function of scientific writing is to communicate knowledge. This viewpoint seems to be popular among faculty. For example, on a

questionnaire given to a convenience sample of eight members of TIMS and ORSA, all agreed that "it helps to write clearly when you submit a paper for publication."

But the rational viewpoint conflicts with the conclusions of some observers. Mahoney [1976; p. 85] gave advice to the researcher who plays the publication game: "Whenever you have a choice between common language and technical argot, use the latter." Authors who ignore this advice to avoid clear writing do so at some personal risk, says Mahoney [1976, p. 961].

Anecdotal evidence is available. For example, I heard of one paper that was rewritten numerous times to improve clarity. It was submitted for publication but was quickly rejected. The author then sent the first draft to the same journal. Although she felt this paper to be incomprehensible, it was accepted for publication.

In my own case, I spent 10 years working on a book [Armstrong, 1978]. Many people advised me that it was a poor use of my time to try to improve the clarity of writing. Furthermore, two faculty members who read early and late versions of the same chapters of my book expressed a preference for the early versions (ones that I thought were poorly written).

Although much anecdotal evidence favors the Dr. Fox hypothesis, I was unable to find empirical evidence with which to compare it with the "communication-for knowledge" hypothesis.

This paper describes two tests of the hypothesis. The first relates the readability of management journals to their prestige. (Dr. Fox says, "more unintelligible journals have higher prestige.") The second test uses concluding sections from academic articles to compare more intelligible and less intelligible presentations of the same information. (Dr. Fox says "competence will be rated more highly for less intelligible passages.")

Readability and Journal Prestige

The communication-for-knowledge hypothesis implies that better journals can attract better authors. They would also devote more effort to refereeing and editorial assistance. Finally, authors would be motivated to do a better job in view of the prestige of the journal. This hypothesis suggests a negative relationship between reading difficulty and prestige.

The Dr. Fox hypothesis was described above. Readers will assume that less intelligible journals are more competent. Thus, a positive relationship should be found between reading difficulty and prestige.

My original study [Armstrong, 1979] examined four fields: management, economics, sociology, and psychology. Different methods were used to rate prestige and different methods were used to estimate unintelligibility. This paper reports only on the examination of the management journals.

Many ways exist to make writing less intelligible. You can use faulty logic. You can convert words to numbers as illustrated in the clever "1 + 1 = 2" paper by Siegfried [1970]. Finally, you can violate rules for clear writing to reduce readability. I examined this last approach.

Readability was estimated for 10 management journals by Loveland et al. [1973]. They used the Flesch Reading Ease Test [Flesch, 1948]. This test is a crude measure of readability because it uses only S, sentence length in words, and N, the number of syllables per 100 words: $F = 207 - 1.02 S - 0.85 N$. Loveland et al. selected twenty 100-word samples at random from each of two years, 1967 and 1971. A close correspondence was found between the reading ease in 1967 and 1971 ($r = 0.92$) In other words, the estimates of reading ease were reliable. To increase reliability, an average score for each journal was computed by combining 1967 and

1971. According to the Spearman-Brown formula [Selltiz et al., 1965, pp. 182183], the reliability of this combined measure should have an $r = 0.96$.

The journals differed substantially in their readability (the coefficient of variation was 30%). This was a favorable factor in the study. Without significant variation it would not have been possible to identify correlates of any type.

I obtained prestige ratings for these 10 management journals by surveying a convenience sample of 20 faculty members. The respondents were primarily from the University of Pennsylvania (11), but also from Drexel University (5), University of Toronto (2), University of Utah (1), and Idaho State (1). A written questionnaire, which asked for a ranking of the academic prestige of each of the journals, was handed to the respondents. The order of presentation of the journals was varied using alphabetical (7), midpoint start (7), and reverse alphabetical order (6). Finally, the wording was varied: four of the questionnaires asked about prestige in terms of scientific standards, while the others did not receive such elaboration. Also, some questionnaires (12) asked for a ranking while others (8) asked for a rating (with 10 being the highest rating).

The results were similar across respondents. For example, the Spearman rank correlation between the 11 University of Pennsylvania rankings and those by 9 other faculty was 0.89. In short, the prestige ratings were reliable. Also, there were substantial differences among journals on the prestige index (coefficient of variation = 35%). Once again, this offered good potential for finding correlates.

According to the Dr. Fox' hypothesis, low scores in readability (difficult reading) should be associated with low scores on the prestige index (high prestige). As shown in Table 1, a positive relationship was found, thus supporting Dr. Fox. The simple correlation of +0.67 was

statistically significant at the .05 level (two-tailed *t* test). However, only 130 of the variance was explained (r^2 adjusted by Lord's formula as recommended by Uhl and Eisenberg [1970]).

An attempt was made to control for other variables that might influence the prestige rating. Measures were obtained for journal circulation and for the number of publications that carry abstracts of papers from that journal. Although these variables were related in the expected way (higher circulation and more abstracting were associated with higher prestige), the relationships were weak and did not add to predictive power.

Table 1. Readability vs. Prestige: Management Journals*

Journal	Prestige	Reading Ease (Flesch)
Administrative Science Quarterly	1.5	20.2
Harvard Business Review	2.2	31.7
Academy of Management Journal	2.5	28.7
California Management Review	2.9	32.6
Industrial Relations	3.3	23.3
Advanced Management Journal	3.6	46.0
Systems & Procedures Journal (New Title: Journal of Systems-Management)	3.7	32.8
Business Horizons	4.5	29.4
Personnel	4.7	35.5
Supervisory Management	5.3	54.3

* Note that prestige decreases as the index goes up, and reading ease increases with increasing Flesch Score.

Equivalent Passages Test

It might be argued that more prestigious journals discuss more difficult issues and this, in turn, might require more difficult language. This would provide an explanation for the results favoring Dr. Fox in the above analysis of journal prestige.

When writing *Long-Range Forecasting* [Armstrong, 1978], I did not find any need to alter content to improve readability. The final version of the book achieved a Gunning Fog Index of 13, which is equal to that for the *Wall Street Journal*, about first year college level, but substantially more difficult than *Love Story*. This suggests that difficult ideas can be written in simple ways. (The Gunning Fog Index (G) is based on average sentence length (S) and the percentage of words (W) with three or more syllables; $G = 0.4(S + W)$. It is designed to approximate the grade level of education needed to understand the material.)

To control for the explanation that difficult ideas require difficult writing, I selected conclusions sections from papers in four management journals: Armstrong [1975], Cort and Dominquez [1977], Kotler and Connor [1977], and Parkan and Warren [1978]. These passages were rewritten to alter the readability index without altering content. I found it possible to simplify the writing without any apparent change in the content. This was accomplished by such steps as:

- (a) eliminating unnecessary words (generally adverbs and adjectives),
- (b) substituting easy for difficult words, and
- (c) breaking long sentences into two sentences.

Additional guidelines were borrowed from Strunk and White [1972]. In addition, by reversing these guidelines I created more difficult versions for two of the studies.

The resulting passages showed a wide variation in readability. For example, the Gunning Index for Armstrong [1975] ranged from 8.3 to 21.8. The more difficult versions were at about the level of journals with high fog indexes. The easy passages were, however, much easier than the easiest journals. Table 2 presents the Gunning Fog Indices for each version.

Table 2. Readability Indices for Equivalent Passages
(original version was “moderate” in each case)

Sources	Version	Gunning Fog Index	Competency Ratings by Faculty
Cort and Dominquez [1977, p. 192]	Easy	9.8	3.2
	Moderate	16.0	3.9
	Difficult	21.6	3.6
Armstrong [1975, p. 116]	Easy	8.3	3.7
	Moderate	15.4	4.3
	Difficult	21.3	4.1
Kotler and Connor [1977, p. 76]	Easy	7.6	3.5
	Moderate	14.0	3.6
Parkan and Warren [1978, p. 119]	Easy	10.2	2.9
	Moderate	16.7	4.6

Subjects were given questionnaires containing one version of each of the four passages. The passages were assembled by using each of the 36 possible combinations. The instructions were:

On the following pages, we have attached samples from papers that have been published in academic journals. The samples represent the conclusions sections from different papers. Please read each sample carefully. On the basis of each sample, please rate the competence of the research that is being reported.

The subjects were not told the names of the journals or authors.

Faculty members in management from Wharton, New York University, and Columbia University were surveyed early in 1979. After numerous follow-ups, replies were received from 32 faculty members. Over 87% of the respondents had acted as referees for journal articles. When asked if they could guess the purpose of this study, 22% did not respond, 50% said no, 12% guessed wrong and 16% guessed right.

The respondents rated competency on a scale ranging from 1 (Highly Incompetent) to 7 (Extremely Competent). They also stated their confidence in the competency rating using a scale running from 1 (Not at All Confident) to 7 (Extremely Confident).

The average competency ratings from each group are provided in the last column of Table 2. The faculty rated the easy versions substantially lower than the other versions. This result was significant at $p < .05$ using the Mann-Whitney Test [Siegel, 1956].

These results add only modest support to the Dr. Fox hypothesis because the sample size was small (only 32 faculty members and only four passages). Furthermore, the study did not directly test whether the faculty members viewed the passages as coming from a source that *they* felt to be legitimate (although this shortcoming might be expected to work against Dr. Fox). On the positive side, the faculty members reported a modest degree of confidence in their ratings (4.3 on the 1 to 7 scale).

Conclusions

A study of 10 management journals found that those more difficult to read were rated higher in prestige by a sample of 20 faculty members. Ratings of easy, moderate, and difficult versions of four otherwise equivalent passages were obtained from 32 faculty members: those passages that were rated easy in readability were judged to be less competent in terms of the research.

Overall, the evidence is consistent with a common suspicion: Clear communication of one's research is not appreciated. Faculty are impressed by less readable articles. As Dr. Fox would say, do not invest energy in writing. Williams and Ware [1976] suggest that lack of clarity is especially helpful when content is poor. As the old saying goes: "If you can't convince them, confuse them."

Improvements in the clarity of academic journals are unlikely, then, to be initiated by researchers. If clarity is a goal for a journal, the editors must take actions, such as:

- Calculate readability indices for each paper that is submitted. (For short papers, the calculation could be based on the complete paper. For longer papers, say over

15 double-spaced pages, one could select a sample of pages. The calculation will take about one hour.)

- Use the readability index in the review process. For example, papers that are difficult to read would be returned for revisions.
- Publish the readability index. The Gunning Fog Index is easiest to interpret. This will help the reader to be aware that a failure to understand may be due to poor writing.*

Such a program would aid in the communication of knowledge. It's cheap. Let's do it. Now I

Acknowledgements

Larry M. Robbins, Director of the Wharton Communication Program, provided useful comments. Mark Pollack assisted in the survey and analyses.

References

- Armstrong, J. Scott, 1979, "Unintelligible Research and Academic Prestige: Further Adventures of Dr. Fox," Working Paper Number 79-015, Department of Marketing, Wharton School, University of Pennsylvania, Philadelphia, Pa. 19104.
- Armstrong, J. Scott, 1975, "Monetary Incentives in Mail Surveys," *Public Opinion Quarterly* 39, pp. 111-116.
- Armstrong, J. Scott, 1978, *Long-Range Forecasting: From Crystal Ball to Computer*. Wiley-Interscience, New York.
- Cort, Stanton G. and Dominguez, Luis V., 1977, "Cross-Shopping and Retail Growth," *Journal of Marketing Research* 14, pp. 187-192.
- Flesch, R., 1948, "A New Readability Yardstick," *Journal of Applied Psychology* 32, pp. 221-235.
- Kotler, Philip and Connor, Richard A., Sr., 1977, "Marketing Professional Services," *Journal of Marketing* 41, pp. 71-76.

- Loveland, John et al., 1973, "An Analysis of the Readability of Selected Management Journals," *Academy of Management Journal* 16, pp. 522-524.
- Mahoney, Michael J., 1976, *Scientist as Subject: The Psychological Imperative*, Ballinger, Cambridge, Mass.
- Naftulin, Donald H., Ware, J.E., Jr., and Donnelly, F. A., 1973, "The Doctor Fox Lecture: A Paradigm of Educational Seduction," *Journal of Medical Education* 48, pp. 630-635.
- Parkan, Celik, and Warren, E. H., Jr., 1978, "Optimal Reneging Decisions in a G/M/1 Queue," *Decision Sciences*, 9, pp. 107-119.
- Powers, R. D., Sumner, W.A., and Kearn, B.E., 1958, "A Recalculation of Four Adult Readability Formulas," *Journal of Educational Psychology* 49, pp. 99-105.
- Selltiz, Claire, Jahoda, M., Deutsch, M., and Cook, S. W., 1965, *Research Methods in Social Relations*. Holt, Rinehart and Winston, New York.
- Siegel, Sidney, 1956, *Nonparametric Statistics*, McGraw-Hill, New York.
- Siegfried, John J., 1970, "A First Lesson in Econometrics," *Journal of Political Economy* 78, pp. 1378-1379.
- Strunk, William, Jr. and White, E. B., 1972, *The Elements of Style*, Macmillan, New York.
- Uhl, Norman and Eisenberg, Terry, 1970, "Predicting Shrinkage in the Multiple Correlation Coefficient," *Educational and Psychological Measurement* 30, pp. 487-489.
- Williams, R. G., and Ware, J. E., 1976, "Validity of Student Ratings of Instruction Under Different Incentive Conditions: A Further Study of the Dr. Fox Effect," *Journal of Educational Psychology* 68, pp. 48-56.