

Readability and Prestige in Scientific Journals

J. Scott Armstrong

The Wharton School, University of Pennsylvania, Philadelphia, PA

Journal of Information Science, 15 (1989) 123-124

© 1989, Elsevier Science Publishers B.V.

Introduction

Hartley, Trueman and Meadows [3] contribute useful evidence on whether scientists can gain prestige by writing in a manner that is difficult to read. This has been called the bafflegab theory by some. They concluded that the evidence was not strong. This comment examines some of the conclusions from Hartley, Trueman, and Meadows (HTM) and recommends directions for further research.

The Hypothesis

I suggest a revised hypothesis for future testing. The revised hypothesis assumes a threshold effect for bafflegab. That is, prestige is related to complexity up to a certain level, beyond which there is no gain. One possible reason is that simple writing is not representative of what readers perceive as scientific writing. Alternatively, it may be that the reader's inability to understand the material will cause, him to attribute his failure to the high competency of the researchers, rather than to his own shortcomings. Furthermore, it may only be necessary to pass a given threshold to achieve this effect.

The threshold hypothesis is consistent with the results from the experimental study reported in Armstrong [2]. When the conclusions were rewritten, the original and the more complex versions did not lead to different results; only the simpler version led raters to judge the writer to be less competent.

The threshold hypothesis will affect the experimental design. If one cannot obtain materials with "simple versions", then little effect would be observed by comparing readability and prestige. The sample of management journals that I analyzed was useful because it contained some journals that were very easy to read. This was not true for my samples in sociology, economics and psychology [1], nor was it true for the psychology journals sampled by Hartley and Trueman as reported in HTM [3].

Analysis

My conclusion from the evidence provided by HTM differs somewhat from HTM's conclusion. Rather than weighting by sample size, I would use an unweighted average of the correlation coefficients. This is because sampling error is probably one of the least important reasons for the differences among the samples. My method yields an average r of -0.24 , compared with the -0.05 for the weighted average. True, this is not a strong effect; however, the measurement problems and the threshold hypothesis could explain why the effect seems small in

these tests. Given the lack of reliability of the readability measure for the economics journals, it may make sense to eliminate that result; in this event, the average r would be -0.39 .

Meta-analysis (the study of studies) can be applied to the prior research. In particular, the tests of statistical significance should incorporate all prior studies. Rosenthal [5] summarizes methods for combining results from independent studies. I applied the "adding t's" method to the five studies reported in HTM's Table 1. The results favored the bafflegab hypothesis and were significant at the 0.05 level using a one-tail test. Interestingly, the results existing prior to HTM were not statistically significant. It is only by considering their evidence that we reach this conclusion.

In addition to the above studies, Armstrong [2] asked subjects to rate the competency of research when presented with four samples of research papers. The content of these papers was the same, but they varied in reading difficulty. The passages that were easy to read were judged as less competent pieces of work ($p < 0.05$, $n = 32$ faculty members). This study was extended by Tracy and Lee [4]; they found that MBA students rated more difficult passages as more competent. Their results were significant at $p < 0.001$. Thus, two quite different approaches yield statistically significant results that are consistent with the bafflegab hypothesis.

Future Research

HTM conclude that it is necessary to develop superior measuring instruments before proceeding with further research. Certainly, better measures would be helpful, but I do not see this as necessary. The use of meta-analysis can help to identify relationships even though the measuring instruments are imprecise. Thus, more studies of the type already reported would be useful. Different approaches would also be useful. These analyses should examine the threshold effect. To do this, some very easily read materials must be included. If only normal materials are used, the threshold effect may not be observed because most of these materials are difficult to read. Given materials of roughly equal readability, other factors would be expected to be important. For example, more prestigious journals may have better copy editors. Also, they may have more important results to report' and thus less need to confuse the reader.

Conclusion

As HTM suggest, better measures would help. But this is unlikely to occur, given the vast effort already devoted to this issue of assessing readability. I believe it is more important to study situations that will allow for further tests of the bafflegab hypothesis. Results from these studies should then be included with the prior studies using meta-analysis in an attempt to determine the conditions under which the bafflegab hypothesis is true. Certainly there must be situations where scientists should write clearly. One possibility is that they should write clearly when they have something important to say.

Acknowledgement

The author thanks J. Hartley for his helpful comments on an earlier version of this paper.

References

J.S. Armstrong, Unintelligible research and academic prestige: further adventures of Dr. Fox, Working paper no 79-015, The Wharton School, University of Pennsylvania Philadelphia, PA.

J.S. Armstrong, Unintelligible management research an academic prestige, *Interfaces* 10 (1982) 80-86.

J. Hartley, M. Trueman, and A.J. Meadows, Readability and prestige in scientific journals, *Journal of Information Science* 14 (1988) 69-75.

Tracy, L.N. and J.A. Lee, Acculturation of graduate business students to academic values: abstrusiveness as a criterion of competence, *Collegiate News and Views* (Fall Winter) (1984) 37-42.

R. Rosenthal, Combining results of independent studies, *Psychological Bulletin* 85 (1978) 185-193.