How to Experiment

Walter F. Tichy
University of Karlsruhe, Germany

Nov. 1997

For the reader eager to learn about the role of experimentation in general, I suggest the following literature. Chalmers, "What is this Thing Called Science" [3] is enligthening about the philosophical underpinnings of the scientific process. It discusses, among others, inductivism, Popper's falsificationism, Kuhn's paradigms, objectivism, and the theory dependence of observation. Latour, "Science in Action" [11] describes the social processes of science-in-the-making as opposed to ready-made science. The fact-building and convincing power of laboratories is illustrated with fascinating examples. I devoured both of these books. I highly recommend reading at least one of them.

The following literature concerns experimentation in computer science. For an appreciation of how software researchers experiment, sample the material cited in the main article and the papers by Basili and Frankl [1, 8]. A good example of a repeated experiment in compiling is Brett [2]. Denning [6] argues that performance evaluation is an excellent form of experimentation in computer science. A landmark in making computer architecture research quantitative is Hennesey's and Patterson's classic book "Computer Architecture" [9]. Cohen, "Empirical Methods in Artificial Intelligence" [5] covers empirical methods in AI, but a large part applies to all of computer science. An excellent discussion of experimental designs as well as a wealth of material on experimentation with software appears in Fenton and Pfleeger, "Software Metrics" [7].

General experimental methods are presented by Christensen, "Experimental Methodology" [4] and Judd at al., "Research Methods in Social Relations" [10]. There are numerous titles regarding statistics; an excellent introductory text is Moore [12]. One of the best stastical packages available today is S-Plus. The book "Modern Applied Statistics with S-PLUS" by Venables and Ripley [13] is both a guide to using S-Plus and a course in modern statistical methods. However, keep in mind that sophisticated statistical analysis is no substitute for good experimental design.

References

[1] Victor R. Basili. The role of experimentation in software engineering: Past, current, and future. In *Proc. of the 18th International Conference on Software Engineering*, pages 442–449. IEEE Computer Society Press, March 1996.

[2] Bevin R. Brett. Comments on "The cost of selective recompilation and environment processing". *ACM Transactions on Software Engineering and Methodology*, 4(2):214–216, 1995.

[3] Alan F. Chalmers. *What is this Thing called Science?* Milton Keynes: The Open University Press, 1988.

[4] Larry B. Christensen. *Experimental Methodology.* Allyn and Bacon, 6th edition, 1994.

[5] Paul R. Cohen. *Empirical Methods for Artificial Intelligence.* MIT Press, 1995.

[6] Peter J. Denning. Performance evaluation: Experimental computer science at its best. *ACM Performance Evaluation Review (SIGMETRICS)*, 10(3):106–109, 1981.

[7] Norman E. Fenton and Shari Lawrence Pfleeger. *Software Metrics: A Rigorous and Practical Approach (2nd ed.).* International Thomson Computer Press, 1997.

[8] Phyllis G. Frankl and Stewart N. Weiss. An experimental comparison of the effectiveness of branch testing and data flow testing. *IEEE Transactions on Software Engineering*, 19(8):774–787, August 1993.

[9] John L. Hennessy and David A. Patterson. *Computer Architecture: A Quantitative Approach.* Morgan Kaufmann Publishers Inc., Santo Mateo, California, 1990. 3rd Printing.

[10] Charles M. Judd, Eliot R. Smith, and Louise H. Kidder. *Research Methods in Social Relations.* Holt, Rinehart and Winston, sixth edition, 1991.

[11] Bruno Latour. *Science in Action: How to follow Scientists and Engineers through Society.* Harvard University Press, 1987.

[12] David S. Moore and George P. McCabe. *Introduction to the Practice of Statistics.* W.H. Freeman and Company, New York, 1993.

[13] William N. Venables and Brian D. Ripley. *Modern Applied Statistics with S-PLUS.* Springer Verlag, 2nd edition, 1997.