American Psychologist
January 1991 Vol. 46, No. 1, 78-79

Some Everyday Thoughts on Ecologically Valid Methods

Mahzarin R. Banaji
Yale University
Robert G. Crowder
Yale University

Judging by the comments in response to our article (Banaji & Crowder, September 1989), enthusiasts for the everyday memory movement are razor sharp in their reading of titles but a little fuzzy on what follows. Thus, many responded as if our article had been (a) an attack on research set in a naturalistic context and (b) insensitive to the application of research findings. It was not. We distinguished two dimensions in scientific research–the ecological validity of the method that is used and the generality of the conclusions permitted. We thought we made clear our view that the best possible situation is for research to be high in both qualities. To repeat, we believe now and plainly stated then that if research is scientifically sound it is better to use ecologically lifelike rather than contrived methods (p. 1188). On some occasions the ability to exert a great deal of control in real-world settings is possible, and in such instances an ingenious experiment, high on both dimensions of ecological validity and scientific generality, is possible. Nobody in his or her right mind would quarrel with this priority. The only issue for debate arises when one dimension must be sacrificed in favor of the other. We argued for the option that guarantees greater generalizability of findings, sacrificing when we must ecologically valid methods of investigation. We wonder whether our critics really endorse the contrary sacrifice.

We are likewise perplexed by the charge that our article "reinforces the academic snobbery that makes application so difficult to achieve" (Gruneberg, Morris, & Sykes, 1991 , this issue, p. 75). The implication of this accusation forces us to respond, again, by reiterating a statement from our original article. We stated that "We do not wish to condone smugness about the generality of laboratory principles to any external context. In fact, we need to test these applications assiduously" (Banaji & Crowder, 1989 , p. 1191).

Nor do we share Gruneberg et al.'s (1991) confusion about what appears to us to be a clear distinction between science and invention. Our discussion was aimed exclusively at methods of scientific inquiry and not at how Edison went about inventing the light bulb or for that matter, the electric chair. We would be delighted to see a Steve Jobs of human memory technology emerge some day, but we are of the belief that our goal is to advance science and not to come up with an occasional memory aid. In addition, we find their distinction between "real" science and "social" science to be alarming because it

expresses the browbeaten timidity of those who regard social science as a less than real science.

Of the several commentaries we received in response to our article (Banaji & Crowder, 1989) we found pieces by Bahrick (1991, this issue) and Gruneberg et al. (1991) to be useful. Bahrick makes an obvious but valid point that we did not mention, that some research questions simply require abandoning the laboratory. For example, a controlled study of forgetting over 25 years might occasionally be feasible, but for forgetting over twice that period, it would be vanishingly unrealistic. He himself has responded thoughtfully by facing head-on some of the challenges in doing field research on such long retention intervals. As much as we admire this work, we must ask: What has been discovered about how memory works? One lesson that Bahrick cites is that forgetting proceeds rapidly at first, and then slows down greatly to reach a nonzero asymptote. Ebbinghaus (1885/1964) demonstrated the same form of forgetting loss in the very first scientific study of memory ever conducted. Peterson and Peterson (1959) verified it in a totally different situation, Underwood (1957) and Sperling (1960) in still others. Most of us thought we already knew the general form of measured forgetting from more than a century of solid research. By analogy, we would admire the technical skill and ingenuity of a physicist who measured specific gravitational effects between two binary stars in a distant galaxy, but we would not feel we understood the principle of gravity any better by virtue of that measurement. We regard attempts to establish the generality of our laboratory findings as important. But we should be unambiguous about when these everyday memory demonstrations are new discoveries and when they are affirmations of scientific principles from the laboratory.

Forty years ago, in a report to the Office of Naval Research, Garner (1950) compared the relative worth of laboratory experimentation versus operational experimentation (i.e., experimental manipulations in the operational field). He concluded that "operational experimentation is more time consuming, far more expensive, and frequently cannot control experimental factors, so that as a practical matter it is very difficult to do operational experimentation which has a high degree of generality of prediction" (p. 2). Garner's point was simple and powerful, and one we assumed was well accepted by the discipline: If you wish to do research that is useful (i.e., practical, functional) the optimal path is controlled experimentation. A nice illustration of this point is Aaronson's (1984) demonstration of the value of ecologically sterile methods in producing direct applications for reading technologies for the visually handicapped. So let us not be confused about the intention of experimental psychologists or the applicability of their findings. On the other hand, in the very best of examples of research sacrificing control for ecological validity, we find that little was learned and at great cost. Bahrick (1991) argues that scientific contributions cannot be evaluated "simply on the basis of meeting arbitrary standards of precision or generalizability" (p. 76). Again, we agree that often a question is important enough that the inability to meet a $p = .05$ standard of reliability, for instance, is not the criterion to use in judging the merit of preliminary research. What we questioned in our article was the seemingly easy abandonment of control in favor of the superficial attraction of ecologically valid methods. We are particularly intrigued by some everyday memory researchers' willingness to forfeit ordinary standards of scientific

protocol because of a perceived failure of the experimental paradigm. What course might the history of other sciences have followed had the experimental method itself been in question each time a scientific problem proved difficult? Would not our critics themselves think it odd if the difficulties encountered in obtaining cold fusion led physicists to become cynical about doing research in the laboratory altogether?

Part of the disenchantment with the experimental method in psychology may stem from an erroneous conception of the discipline of psychological science. This conception, not peculiar to everyday memory researchers, emerges from the misguided notion that humans (as the unit of analysis of most psychological research) are infinitely more complex, mysterious, variable in behavior, and resistant to understanding than other entities in the universe. The supposedly greater complexity of human behavior necessitates leaving controlled environments that apparently suffice in studying less complex phenomenon such as light, time, gravity, and entropy! It is such misinformation that perhaps led Aanstoos (1991, this issue) to the howler that "if two chemicals are going to react, they'll react when combined in a test tube just as well as outside it" (p. 77). It is precisely because this is not true that test tubes were invented in the first place.

The notion that human behavior is vastly more complex than chemical reactions can only emerge from a lack of acquaintance with other sciences. The random behavior of the photon in the two-slit experiment makes it impossible to determine when it will behave as a particle and when it will behave as a wave. Does Aanstoos really wish to claim that such fundamental uncertainty in nature is less complex, more straightforward, and therefore more amenable to laboratory study than the memory phenomena we study?

In conclusion, some warm commentary about our article was not entirely unexpected. Bankruptcy may not have been the choicest term, but readers should be aware that we considered and rejected some even more provocative ones. We note that the reactions derive from inattention to the article itself (as opposed to the title), confusion between science and technology, and naivete about the former.

References

Aanstoos, C. M. (1991). Experimental psychology and the challenge of real life. *American Psychologist, 46*, 77-78.

Aaronson, D. (1984). Computer methods and ecological validity in reading research. *Behavior Research Methods, Instruments, & Computers, 16*, 102-108.

Bahrick, H. P. (1991). A speedy recovery from bankruptcy for ecological memory research. *American Psychologist, 46*, 76-77.

Banaji, M. R. & Crowder, R. C. (1989). The bankruptcy of everyday memory. *American Psychologist, 44*, 1185-1193.

Ebbinghaus, H. E. (1964). *Memory: A contribution to experimental psychology.* (New York: Dover. (Original work published 1885)

Garner, W. R. (1950). The validity of prediction from laboratory experiments to naval operational situations in the area of human engineering and systems research (Report No.

166-I-130).(Baltimore: Johns Hopkins University, Institute for Cooperative Research)
Gruneberg, M. M., Morris, P. E. & Sykes, R. N. (1991). The obituary on everyday memory and its practical applications is premature. *American Psychologist, 46*, 74-75.
Peterson, L. R. & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74* , (Whole No. 498).
Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64*, 49-60.