They Read an Article?

A Commentary on the Everyday Memory Controversy

Henry L. Roediger III *Rice University*

In a provocative essay, Neisser (1978) argued that memory research conducted under controlled laboratory conditions has produced no important knowledge and that investigators should turn to uncontrolled, naturalistic observation outside the lab. Banaji and Crowder (1989) argued for the importance of experimental control in producing generalizable knowledge. The authors of the current target articles attack Banaji and Crowder, but they never confront the main issue: whether experimental control is necessary or desirable in research. The commentators show that controlled research can be done outside the lab, but such examples reinforce Banaji and Crowder's main point about the necessity of experimental control. Naturalistic observation is a good starting point for research, but a poor substitution for controlled experimentation to uncover the important factors at work.

In 1954 Hastorf and Cantril published "They Saw a Game: A Case Study," in which they reported an experiment capitalizing on a naturally occurring rivalry to examine its influence on perception. Today we might want to place their research in an everyday perception tradition, but they placed it in the framework of transactional analysis (e.g., Kilpatrick, 1952). The event motivating their article was an acrimonious football game in which Dartmouth played an undefeated Princeton team. The game was marred by fights, penalties, and several injured players. Among the injured was Princeton's star, Dick Kazmaier, who had just been featured on the cover of *Time* magazine. Princeton won the game, but debate continued in the U.S. press for several weeks about which team was at fault for the game's roughness. Unsurprisingly, students at each school blamed the opponent team in their school newspapers and in interviews.

The polarized student opinion led to Hastorf and Cantril's (1954) experiment. They had groups of students at each school watch a film of the game and judge, among other things, the number of infractions committed by each team during the game. The Dartmouth students counted 4.3 infractions committed by their own team and 4.4 by the opponents, reflecting the general Dartmouth belief that the game had been rough but fair (and contrary to accusations that Dartmouth's Big Green team had unfairly attacked the Tigers' star). On the other hand, the Princeton students judged the Dartmouth team to have committed 9.8 infractions and their own team only 4.2, confirming their claim that the Dartmouth team had resorted to dirty tactics to try to stay even with Princeton, which was "obviously the better team," according to a writer for the *Daily Princetonian.*

I bother to recount this study because the brouhaha about Banaji and Crowder's (1989) article seems to recapitulate the same psychological principles as those revealed in Hastorf and Cantril's (1954) study. It could almost be seen as a conceptual replication. In my judgment, the main commentators on the target article of this inquiry (i.e., Conway, 1991, this issue; Ceci & Bronfenbrenner, 1991, this issue) seem remarkably able to perceive offense where none was intended, but more curiously still, to miss, ignore, or duck the main point raised by Banaji and Crowder. At some point in reading the commentaries, it occurred to me that I was actually participating in a clever experiment by Elizabeth Loftus, the editor of this Science Watch section, to see if my memory for the Banaji and Crowder article could be systematically distorted by the commentators' remarks, because often the points they rebutted bore little resemblance to my memory of those in the original article. (No follow-up questionnaire has yet appeared to test me on Banaji and Crowder, so I must reject this hypothesis as to the source of the misinformation.) However, this was one of those cases in which the misinformation was so blatant as to not be credible, as can be verified by rereading Banaji and Crowder's original article, as I did. Of course, I am no less immune to the sorts of misperception noted by Hastorf and Cantril (1954) than is anyone else and so will provide my own view of the matter here.

The starting point for this little controversy, in the eyes of the commentators, is Banaji and Crowder's (1989) article. But Banaji and Crowder could justifiably claim, in line with the Billy Joel song, that "We didn't start the fire." It was already there, if not since the world has been turning, at least since it was kindled by Neisser's (1978) inflammatory essay in which he attacked the laboratory tradition of memory research. This essay has been taken as the rallying cry of everyday memory researchers, at least until the appearance of Banaji and Crowder's article. For example, in her textbook *Memory in the Real World,* Cohen (1989) discussed Neisser's essay (and the talk on which it was based) this way:

> In 1976, at a conference on Practical Aspects of Memory, Ulric Neisser gave a talk (later published in 1978) entitled "Memory: What Are the Important Questions?" in which he dismissed the work of the past 100 years as largely worthless. This talk was undoubtedly a milestone in the psychology of memory. Neisser believes that the important questions about memory

Vol. 46, No. 1, 37–40

are those that arise out of everyday experience. . . . The traditional laboratory experiments, according to Neisser, have failed to study all the most interesting and significant problems and have shed no light on them. He claimed that the experimental findings are trivial, pointless, or obvious, and fail to generalize outside the laboratory. (1989, p. 2)[1]

Cohen went on to remark that "Neisser's ideas had an enthusiastic reception," (p. 2) at least in some circles (but see Hintzman, in press). As I have noted elsewhere (Roediger, 1990), Neisser's dismissal of memory research did not display much acquaintance with the topic, because many important topics have long been studied in laboratory contexts, as examination of Baddeley's (1976) and Crowder's (1976) texts, based largely on laboratory work, will bear out.

Banaji and Crowder's (1989) article was therefore not a bolt from the blue, attacking a new research tradition, but a reply to Neisser's (1978) argument that researchers should dismiss all laboratory work and look only to ecologically valid situations lacking in experimental control. I see their essay as arguing primarily for the necessity of experimental control in drawing valid, potentially generalizable conclusions. In their two-by-two array of approaches, shown on page 1188 of their article, they argued that high ecological validity and high generalizability is the best research situation, but that if one dimension must be sacrificed it should be ecological validity and not generalizability. Contrived laboratory methods are favored in establishing solid effects of the variables operating in the situation, which leads to generalizability of the results. Put another way, a finding must have internal validity—the observed effect must be securely tied to the independent variable or variables—before a researcher can even hope to generalize to a new situation, whether inside or outside the lab.

As I see it, then, the critical difference between Neisser's (1978) and Banaji and Crowder's (1989) viewpoints is on this issue of experimental control. For Neisser, it was not really needed—psychologists interested in memory need only take off their laboratory blinders and head outside to look at how memory operates in the real world. For Banaji and Crowder (1989), experimental control over the phenomena of interest is paramount in gaining an understanding of those phenomena. Curiously, the commentators defending the everyday memory perspective seem to agree more with Banaji and Crowder than with Neisser on this central issue. No one defended the propositions that experimental control is useless, or that all research in the lab is (and has been) a waste of time because the memory laboratory is not an ecologically valid situation, and so on. Rather, they attempted to show that research done outside the confines of a laboratory can, with ingenuity, sometimes be reasonably well controlled and can lead to firm conclusions. That is fine; I suspect Banaji and Crowder would welcome such research. But they would do so because experimental control has been introduced to provide generalizable conclusions, and not because of the setting of the research. Thus the version of everyday memory research most of the commentators defend seems to be more akin to the laboratory tradition, with its experimental control, than to the uncontrolled, ethological tradition extolled by Neisser (1978). It is worth noting that Neisser (1988) has also moderated his views of a decade earlier.

Assuming that the foregoing remarks are moderately accurate, then many of the commentators complaints about Banaji and Crowder's (1989) article seem misguided, as a rereading of the target article (or even of pp. 1187–1188) would show. Banaji and Crowder did not argue against applied research, as charged by Gruneberg, Morris, and Sykes (1991, this issue), but against uncontrolled applied research from which no generalizable conclusions are possible; similarly, with the same proviso, they had no quarrel with research conducted on special populations or in settings outside the lab, as charged by Ceci and Bronfenbrenner (1991, this issue); certainly they never equated science with laboratory experiments, as was also maintained by Ceci and Bronfenbrenner. And so on. As with the Princeton students reviewing the Princeton–Dartmouth game, the imagined offenses (if so they are) appear quite real to the participants; however, from my perspective, most are not to be found in the article itself.

The commentators point to many examples of interesting studies claimed under the banner of everyday memory research. I agree that most of these exemplify excellent research, but must note that including them under the rubric of everyday memory research (as defined by Neisser's, 1978, essay) is greatly stretching matters. For example, research on the tip-of-the-tongue phenomenon (Brown & McNeill, 1966), the feeling-of-knowing experience (Hart, 1965; Nelson, 1988), eyewitness testimony (Loftus & Palmer, 1974), and reality monitoring (Johnson & Raye, 1981) all largely involved laboratory settings with college students as subjects and, in most cases, rather intrusive techniques to establish experimental control. None of these used the ethological approach advocated by Neisser (1978).

Another example, endorsed by both Ceci and Bronfenbrenner (1991) and Conway (1991, this issue) as demonstrating an important finding from the everyday memory research tradition, is the long-term recency effect found in rugby players' recall of their matches, as reported by Baddeley and Hitch (1977). As Ceci and Bronfenbrenner take Banaji and Crowder (1989) to task for not citing some important and relevant literature (viz., Ceci & Bronfenbrenner's own work), it is worth noting that a 1976 report also in the everyday memory tradition made

[1] Cohen (1989) listed the date of the conference as 1976, but Neisser (1982, p. 3) stated that it occurred in 1978. An opportunity may exist here for naturalistic research to uncover the source of the error and reasons for it, at least within the limits of these types of methods.

I thank Michael J. Watkins for commenting on an earlier version of the article.

Correspondence concerning this article should be addressed to Henry L. Roediger III, Department of Psychology, Rice University, Houston, TX 77251-1892.

the same point, but was somehow missed by Ceci and Bronfenbrenner and by Conway in their discussions. The study, by Roediger and Crowder (1976; reprinted in Neisser, 1982) had students attempt to recall the presidents in order and showed a strong long-term recency effect, as well as a primacy effect and a sort of distinctiveness (or von Restorff) effect in the elevated recall of Lincoln. The paper was rejected by three editors because we did not control the stimulus materials, putting it squarely in the everyday memory tradition, even if we did not know it at the time. The impact of both our study and that of Baddeley and Hitch was lessened by the fact that Tzeng (1973) and Bjork and Whitten (1974) had demonstrated long term-recency in a laboratory situation. All of these reports and others (e.g., Watkins & Peynircioglu, 1983) have been taken as more consistent with a distinctiveness-of-endpoints interpretation of the serial position effect than with the traditional two-store account (but see Raaijmakers, in press).

In my opinion, the characterization of laboratory-based research by several writers is off the mark, but this is not the place to provide a point-by-point rebuttal. However, one major claim by Conway (1991) is so egregiously erroneous that it must be corrected. I refer to his statements to the effect that laboratory researchers have not studied meaningful, knowledge-based processing and that the study of episodic memory in practice is, in the main, a "knowledge-restricted or meaning-restricted approach to the psychology of memory" (p. xx). Surely nothing could be further from the truth: Beside levels of processing research, which he mentions, there are dozens (and maybe hundreds) of studies of semantic elaboration, of semantic priming, of the effect of category clustering based on meaning, of subjective organization, of semantic confusion errors, of semantic false recognition errors, of semantic congruity effects in cued recall, of memory for gist and not surface information, of the benefits of meaningful hierarchical organization, and for that matter, even of the meaningfulness of nonsense syllables. Indeed, some have even argued that laboratory-based research has too strongly emphasized meaning and sought to refute such semantic primacy by showing that other aspects of information were important for memory, too (Kolers & Roediger, 1984, pp. 430–432). The laboratory approach to memory may be guilty of some sins, but ignoring meaning or the prior knowledge of the rememberer is clearly not one of them.

The kind of error Conway (1991) made is most likely encouraged by Neisser's (1978) arguments that the first 100 years of laboratory experimentation were a waste of time. Why bother learning anything about this research tradition if it is all irrelevant? Such historical amnesia leads to some curious developments. Researchers study flashbulb memories without relating them to laboratory studies showing that distinctive events are well remembered; others study the tip-of-the-tongue effect without considering related studies of inhibitory effects of semantic primes; and some study the interfering effects of misleading interpolated information after an event without considering this as an instance of retroactive interference. The last case is particularly interesting because current debate in the eyewitness testimony arena revolves around whether misleading postevent information blends or overwrites memory for the original event (e.g., Loftus, 1979; Loftus & Palmer, 1974) or whether memories for both the original and interpolated events are held simultaneously and the conditions of the test determine which is retrieved and how (e.g., McCloskey & Zaragoza, 1985). To some, this debate sounds reminiscent of two-factor interference theory in which one factor was unlearning the original event during retroactive interference, corresponding to Loftus's position, and the other was response competition that could be affected by test conditions, as in McCloskey and Zaragoza's position. Recent developments in this realm seem to echo those of 30 or so years ago, confirming George Santayana's famous dictum. But, of course, Neisser (1982) encouraged this by arguing that "the experiments of interference theorists seem like empty exercises to most of us. Were they ever anything else?" (p. 9). My answer: Yes, they were fundamentally important, as confirmed by observations of Loftus and Palmer and many others, both inside and out of the lab. Interference is one of the most potent variables affecting memory, as has been shown many times (e.g., McGeoch, 1932; Underwood, 1957).

In my opinion, the everyday memory movement, if removed of some excesses, has much to offer. The traditional role of naturalistic observation is to draw attention to significant phenomena and to suggest interesting ideas. Researchers will then typically create a laboratory analog of the natural situation in which potentially relevant variables can be brought under control and studied. Of course, the danger exists in arranging the laboratory situation (or, equivalently, introducing control into the field) that the critical variables may vanish, as will the phenomenon. But this is not the fault of laboratory research or of the quest for experimental control; it simply means the experimenter failed to select the right variables in bringing the phenomenon into the lab, or in introducing control into the field.

Jenkins (1979) framed the issue of generalizability in a useful way in his tetrahedral model of memory experiments. He noted that any experiment can be considered in the context of four sets of variables: (a) types of subjects, (b) types of materials to be tested, (c) the type of test used, and (d) the research setting (including orienting activities, instructions, subject strategies, etc.). For any particular finding, we can ask whether it generalizes across the other dimensions in Jenkins's tetrahedron, or indeed, across the same dimension when manipulated a different way. But before questions of generality should be asked empirically, one must have an unambiguous finding to generalize. This point leads us back to Banaji and Crowder (1989): Without experimental control over observations, no possibility of generalizability (external validity) exists (see also Mook, 1983). As Donald Broadbent (1990), perhaps our leading applied experimental psychologist, has recently noted in this context, "The last thing an applied psychologist should do is sloppy experiments outside the laboratory." Everyday memory research of the sort Neisser (1978) advocated may alert psychologists to new phenomena, but progress in understanding them will depend on what experimental control is brought to bear, whether inside or outside the laboratory.

REFERENCES

Baddeley, A. D. (1976). *The psychology of memory.* New York: Basic Books.

Baddeley, A. D., & Hitch, G. (1977). Recency reexamined. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 647–667) Hillsdale, NJ: Erlbaum.

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

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173–189.

Broadbent, D. (1990, July). *General discussion of the conference.* Paper presented at the meeting of Attention and Performance XIV: A Silver Jubilee, Ann Arbor, MI.

Brown, R., & McNeill, D. (1966). The tip-of-the-tongue phenomenon. *Journal of Verbal Learning & Verbal Behavior, 5,* 325–337.

Ceci, S. J., & Bronfenbrenner, U. (1991). On the demise of everyday memory: "The rumors of my death are much exaggerated." *American Psychologist, 46,* 27–31.

Cohen, G. (1989). *Memory in the real world.* Hillsdale, NJ: Erlbaum.

Conway, M. A. (1991). In defense of everyday memory: A reply to Banaji and Crowder. *American Psychologist, 46,* 19–26.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Gruneberg, M. M., Morris, P. E., & Sykes, R. N. (1991). The obituary on everyday memory and its practical application is premature. *American Psychologist, 46,* 74–75.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56,* 208–216.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129–134.

Hintzman, D. L. (in press). 25 years of learning and memory: Was the cognitive revolution a mistake? In D. E. Meyer, & S. Kornblum (Eds.), *Attention and performance XIV: A silver jubilee.* Hillsdale, NJ: Erlbaum.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429–446). Hillsdale, NJ: Erlbaum.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67–85.

Kilpatrick, F. P. (Ed.). (1952). *Human behavior from the transactional point of view.* Hanover, NH: Institute for Associated Research.

Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning & Verbal Behavior, 23,* 425–449.

Loftus, E. F. (1979). The malleability of human memory. *American Scientist, 67,* 312–320.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior, 13,* 585–589.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114,* 1–16.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39,* 352–370.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387.

Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). San Diego, CA: Academic Press.

Neisser, U. (1982). Memory: What are the important questions? In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 3–19). New York: Freeman. (Original work published 1978)

Neisser, U. (1988). New vistas in the study of memory. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 1–10). Cambridge, England: Cambridge University Press.

Nelson, T. O. (1988). Predictive accuracy of feeling of knowing across different criterion tasks and across different subject populations and individuals. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory Vol. 2.* (pp. 190–196). New York: Wiley.

Raaijmakers, J. G. W. (in press). The story of the two-store model. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: A silver jubilee.* Hillsdale, NJ: Erlbaum.

Roediger, H. L. (1990). [Review of *Remembering reconsidered: Ecological and traditional approaches to the study of memory*]. *American Journal of Psychology, 103,* 403–409.

Roediger, H. L., & Crowder, R. G. (1976). A serial position effect in recall of United States Presidents. *Bulletin of the Psychonomic Society, 8,* 275–278.

Tzeng, O. J. L. (1973). Positive recency effect in delayed free recall, *Journal of Verbal Learning & Verbal Behavior, 12,* 436–439.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49–60.

Watkins, M. J., & Peynircioglu, Z. F. (1983). Three recency effects at the same time. *Journal of Verbal Learning & Verbal Behavior, 22,* 375–384.