The Bankruptcy of Everyday Memory

Mahzarin R. Banaji and Robert G. Crowder
Yale University

ABSTRACT: A new approach to the study of memory has emerged recently, characterized by a preoccupation with natural settings and with the immediate applicability of research findings. In contrast, the laboratory study of memory relies on experimental techniques for theory testing and is concerned with the discovery of generalizable principles. Although both approaches share the goal of generalizability, they differ sharply in the evaluation of how that goal is best accomplished. In this article, we criticize the everyday memory approach, arguing that ecologically valid methods do not ensure generalizibility of findings. We discuss studies high in ecological validity of method but low in generalizibility, and others low in ecological validity of method but high in generalizibility. We solidly endorse the latter approach, believing that an obsession with ecological validity of method can compromise genuine accomplishments.

Once upon a time, when chemistry was young, questions of ecological validity were earnestly raised by well-respected chemists and were debated at scientific meetings and in scholarly journals. We understand from a colleague (who is a distinguished historian of science but modestly asked not to be named) that partisans of one point of view called themselves the "everyday chemistry movement." They pointed out that the world offered many vivid examples of chemical principles at work in our daily lives—the rising of pastry dough, the curdling of sauces (the great chef Brillat-Savarin was then laying the foundation for the principles of applied chemistry thereafter called French cuisine), the smelting of metal alloys, the rusting of armor, and the combustion of gunpowder. Why not, they asked, study chemical principles in these ecologically faithful settings rather than in tiresome laboratories with their unnatural test tubes, burners, and finicky rules of measurement? The normal world around us, they said, has no end of interesting and virtually unstudied manifestations of chemistry. One scholar, who was famous for his contributions to the new science, even commented that he thought one thing was certain: "If X is an interesting or socially significant aspect of chemistry, then chemists have hardly ever studied X." (Some advocates were actually abusive in their statements; we cite one of the nicer ones).

Of course this parable is apocryphal. Its purpose is to make the point that the other sciences would have been hopelessly paralyzed if they had been deprived of the methods of science during their evolution. Imagine astronomy being conducted with only the naked eye, biology without tissue cultures, physics without vacuums, or chemistry without test tubes! The everyday world is full of principles from these sciences in action, but do we really think their data bases should have been those everyday applications? Of course not. Should the psychology of memory be any different? We think not.

There has been more than a decade of passionate rhetoric claiming that important questions about memory could be tackled if only researchers looked to the "real world" for hypothesis validation. Yet, no delivery has been made on these claims: No theories that have unprecedented explanatory power have been produced; no new principles of memory have been discovered; and no methods of data collection have been developed that add sophistication or precision. In this article, we argue that the movement to develop an ecologically valid psychology of memory has proven itself largely bankrupt and, moreover, that it carries the potential danger of compromising genuine accomplishments of our young endeavor.

Selected papers presented at two conferences on practical aspects of memory (Gruneberg, Morris, & Sykes, 1978, 1988) and other research on autobiographical memory were used as representative work on memory for everyday events. First, however we should define more exactly the targets of our critique and, here, matters of terminology and special populations deserve mention but little more.

Preliminary Issues

Terminology

The term *autobiographical memory* has frequently appeared in recent years as a description of memory for events that have occurred in the "real" life of the subject rather than in the laboratory. The term is misleading because virtually every study conducted on episodic memory since, and including specifically, those of Ebbinghaus (1885/1964), has concerned memory for personally experienced, and therefore autobiographical, information. Autobiographical memory, then, appears to be a term of temporary convenience that has an unintentional denotation and a mysterious connotation.

In a commentary on these issues, Bruce (1985) has suggested the alternative term *ecological memory,* perhaps in response to some of the obvious shortcomings of the terms *autobiographical* and even *everyday* (see also, Sherry & Schachter, 1987). He endorsed a biological–evolutionary approach to memory, rejecting some of the more superficial abuses that have been aimed at the 100-year history of scientific investigations of memory. In

 0003-066X/89/$00.75
Vol. 44, No. 9, 1185-1193

proposing as a model the field of animal behavior, he suggested that psychologists pay more attention to the functional significance of memory in the evolutionary sense. We appreciate his use of this term as constructive and well-motivated. However, we think the biological model is a poor one for how memory functions in humans. The rate of change for systems conducive to the survival and breeding of the species is almost incalculably slow even in the context of human history. If we accept adaptations through *learning* as being the model, of course, we have then accepted the premise of pioneers such as Pavlov, Thorndike, and Hull, whose reductionism was based exactly on an ecological perspective. In this article, we use the term *everyday memory,* which we believe better captures the essential meaning of the new body of research.

Special Populations

A more substantial issue concerns what kinds of research appear to qualify for membership in the everyday memory movement. We discern two kinds of activity from our review of the literature, the first of which we discuss in this section.

Sometimes, the choice of ecologically representative stimulus materials, realistically manipulated states of consciousness, or colorful subject populations seems automatically to confer everyday memory status on an investigation, however scrupulously it otherwise adheres to conventional scientific methods. The studies that fall into these categories are ones that use stimuli or situations that are likely to occur in everyday life, studies of alcohol or drug states and memory, and studies in which the subjects are selected from a deviant population (where deviance is defined as not being a college undergraduate).

The criteria, however, are disjunctive. A research program could, for example, use the most artificial of stimulus populations and the most contrived of tasks, but as long as it studied the effects of sleep deprivation, underwater diving, or alcohol intoxication, it would still be seen as representative of the movement (see sections on neurological memory deficit, drugs and memory, reading, dyslexia, motoric memory, etc., in Gruneberg, Morris, & Sykes, 1988, Vol. 2).

We were bemused by the logic here that a study of paired-associate learning of nonsense syllables would qualify if one group of subjects were Korsakoff patients or idiot savants, but we nonetheless have no quarrel to pick with these studies. In fact, many of them are good examples of the value of the scientific method. Our claim is not that special populations of subjects are always irrelevant to scientific advancement. Far from it—they are often vital to the confirmation of critical aspects of theories. For example, a study of coding short-term memory for deaf subjects (Conrad, 1972) was a legitimate, even urgent, priority given the reliance on phonetic coding established earlier by Conrad (1964) for normal populations. In a similar way, congenitally blind subjects have been shown to perform much in the same way as normal subjects in several "visual imagery tasks" (Kerr, 1983), and this fact constrains theories of how imagery operates in information processing.

More insidious is the second type of activity in which studies qualify for inclusion because the entire phenomenon is defined by the method of naturalistic observation. Into this category would fall retrospective analyses of eyewitness testimony, studies of memory for classroom lectures, oral examinations, or events from a summer vacation. These approaches largely abandon the opportunities for analytic control of the learning and test situations, and it is to these that we direct our skepticism in this article. But first, to place the controversy in historical context, we discuss an influential challenge that was posed to the experimental science of memory.

A Challenge to the Experimental Science of Memory

We proceed by presenting some issues raised by Neisser (1978) in a chapter that is widely cited as the vision of an ecological approach to the study of memory, assuming for now that such a goal is possible in psychology, unlike astronomy, for example. (We take it to be beyond question, here and elsewhere, that we do not mean to personalize the controversy by citing Neisser repeatedly. Some of his pronouncements, besides being influential, are especially articulate and thus are inviting pegs on which to hang our arguments. Indeed, his most recent statement [Neisser, 1988] is considerably less severe in its indictment of laboratory techniques in the study of memory).

The Thundering Silence

The first conference on practical aspects of memory (Gruneberg et al., 1978) began with a talk titled "Memory: What Are the Important Questions?" in which Neisser rebuked psychology's "thundering silence" about questions of great interest, such as how one remembers sources of information, arguments, or material that is relevant to one's current thought. He also pointed to the embarrassment of discovering that we psychologists have no answers for our layperson friend who eagerly poses innocent questions such as the following: Why are there limitations on memories for early childhood? Why is it difficult to remember appointments? Why is it easy to find one's way around one's hometown after a 30-year absence? Why did I forget what I had for breakfast this morning?

We thank Bob Abelson, R. Bhaskar, Denise Dellarosa, Matt Erdelyi, Tony Greenwald, Beth Loftus, Mike Lynn, Jack Mayer, Roddy Roediger, Peter Salovey, and the Memory Lunch Group at Yale for valuable comments and discussions. Inclusion in this list does not necessarily imply agreement with our position.

Preparation of this article was supported in part by a grant from the Social Science Faculty Research Fund at Yale University to M. Banaji and by Grant GB 8608344 from the National Science Foundation to R. G. Crowder.

Correspondence concerning this article should be addressed to Mahzarin R. Banaji or Robert G. Crowder, Department of Psychology, Yale University, P.O. Box 11A, Yale Station, New Haven, CT 06520-7447.

Neisser's (1978) lament was that psychologists are not interested in such questions because they do not believe them to be truly important. This allegation has at least two answers: First, it is a misrepresentation. Psychologists have no delusions that laboratory techniques are their own justification. Rather, many of us believe that the way in which questions about memory can best be answered is through the empirical discovery of facts about memory that have *generalizability,* and not by the use of tasks that carry an illusion of ecological validity by testing memory in everyday contexts. By analogy, our apocryphal chemist might well retain an interest in why cake dough rises during baking but decide that controlled experimentation on yeast or the reactions of moist baking powder to heat would pay off more than loitering in professional bakeries and taking careful notes. (This issue will receive fuller discussion in a later section.)

Second, no embarrassment is in order when a psychologist is confronted with a layperson asking so-called interesting questions about memory. Science is an acquired taste, and scientific priorities may or may not continue to respect the mundane definition of what is "interesting." What other science, we ask, has established that its students should decide on the importance of questions by checking first with Aunt Martha or the expressway toll-taker? Why, and with what value, should the science of memory be singled out among the other sciences and burdened with this absurd criterion of legitimacy? If one wished to maintain that psychology has an inherently different responsibility from those of the other sciences, namely, the responsibility to provide the everyday public with everyday explanations, then one would need to explain why this peculiar demand is attached to psychology. That philosophical analysis is missing from the literature in our judgment, but to refute it here would take us far afield of our agenda.

Counterintuition as a Criterion of Good Science

Another issue in research on memory concerns the findings themselves. Neisser pointed out that enduring principles of memory, such as the effects of meaningfulness, practice, savings at relearning, and so on, are painfully obvious to students (and even to kindergartners!). According to Neisser, this should be yet another embarrassment to psychologists, who discover that the pinnacle of 100 years of slaving in the laboratory is a string of simplistic, intuitive effects. Again, there are several answers to the accusation.

It is our experience that students in introductory courses are often surprised and intrigued when they are introduced to experimental findings about memory, one example being the serial position curve. In fact, one of us has routinely asked her introductory psychology class a question before conducting the well-known classroom demonstration: "I am going to read a list of words to you, such as Apple, Mug, Square, etc. At the end of reading that list, I will ask you to write down as many of the words as you can remember. Before I do that, however, can you tell me which of these words you think you will remember?" Of the many and interesting hypotheses students have generated, rarely has one borne resemblance to the correct answer. Contrary to Neisser's claim, students do not always know these findings before they hear our lectures.

However, even if laypersons do find out that our experimental data only reaffirm their preconceived theories about how memory works, that confirmation should not be a source of embarrassment to us as Neisser has proclaimed. The belief that objects that are thrown up will fall down also corresponds to intuition and everyday observation. Needless to say, if the principle of counterintuition were applied to decisions of scientific worth, Isaac Newton might easily have ignored inventing the calculus. Risking the embarrassment of stating the obvious and intuitive, we say that the question to the scientist is not only that an effect occurs, but why it occurs. That a wise undergraduate can predict that a recency effect will be obtained unless the subject is assigned to an immediate distractor condition cannot belittle the efforts of a scientist interested in the nature of short-term memory.

Myths About Memory

Our students and laypersons in general "know" many things about memory that are complete nonsense. One is that slow learners show less forgetting than faster learners (Underwood, 1964). Another is that rote repetition increases the probability of later recall (Craik & Watkins, 1973; Rundus, 1977). A third is that some lucky adults have photographic memories (see Klatzky, 1984, Chap. 6, for other examples of "commonsense" principles of memory that are just silly in light of evidence that we have). As with intuitive physics (McCloskey, 1983), a systematic body of knowledge is needed for people to sort out which of their many beliefs are worth holding on to and which are worthless.

Intuitive psychology, below the surface, is just as fraught with ignorance as intuitive physics (McCloskey, 1983). If the growth of memory performance with repeated practice is a boring, Ebbinghausian platitude, painfully obvious to the laity, then so is the growth of recall under conditions of maintenance rehearsal. The latter belief, however, is dead wrong. Folk wisdom embraces many correct intuitions, but it also embraces many ideas that are utter nonsense. Our great grandparents knew for sure that mushrooms were poisonous, but they also "knew" that tomatoes were poisonous. A systematic body of knowledge about memory needs to be accumulated in order to separate the myths from the facts, and our experimental techniques will serve well to accomplish this goal.

A Two-by-Two Array of Approaches

The attitude reflected in Neisser's (1978) commentary is based on at least one fundamental confusion, that the use of lifelike methods guarantees generality of conclusions to real-life situations. We argue that ecological validity of the methods as such is unimportant and can even work against generalizability.

This theme can be clarified by the construction of a two-dimensional array of scientific approaches. One dimension is the ecological validity of the methodology. The other is the external generalizability of the conclusions permitted by the research. For simplicity, these may be imagined as a two-by-two array (Figure 1), although in reality we think of them as dimensions. Now, nobody would deny that, other things being equal, the cell in which ecologically valid methods are used to achieve generalizable results is the best situation in which to find oneself. Nor could it possibly be denied that the combination of contrived, artificial methods and conclusions with no external validity produces a sorry state. The only real debates focus on the other two cells, where a cost–benefit analysis must be applied, and it is these two cells that we scrutinize in the rest of the article.

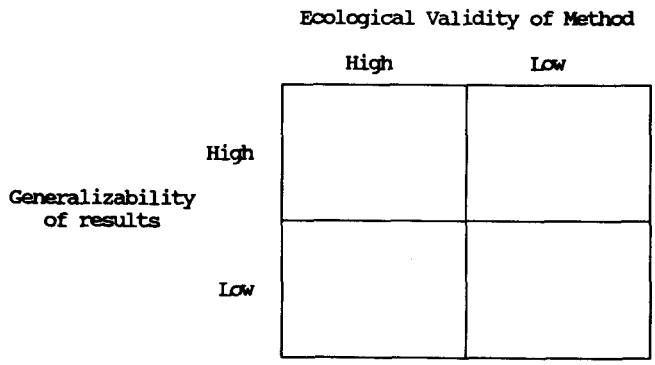
We come down solidly in favor of accepting contrived methods as long as the payoff in generality of conclusions is great enough. Gathering from a survey of the sources cited, we conclude that others of our colleagues in the study of memory would opt for the other cell, lifelike methods at all costs. That strategy, we fear, would lead the psychology of memory into the same stultification as studying backyard astronomy with the naked eye, chemistry in the kitchen, and biology with a walk through the forest. We have nothing against backyards, kitchens, or forests, but they are not ideal settings for the practice of science, and neither is everyday memory. We question whether the principles of learning discovered in the animal laboratory (see Rescorla, 1988) would have emerged if behaviorists had been dedicated to following rats around their natural urban habitats, craning their necks peering into upturned manhole covers.

Of course, some sciences absolutely require naturalistic observation and description in order to define the phenomena under study. Certain areas of ethology, and perhaps primate social behavior, are good examples of areas that depend on naturalistic observation, but we deny that a case has been made for this approach in the study of memory. The method of naturalistic observation can succeed in a science that has developed precise techniques for translating observations into a formal language such that the operations of invariant mechanisms can be shown obviously. For example, even before the theory of problem solving embodied in General Problem Solver (Newell & Simon, 1972) was developed, Duncker (1945) had used verbal protocol data to study problem solving. The difference, besides the availability of tape-recording equipment, was that Newell and Simon were able to cast their "naturalistic" observations into a formal language (i.e., IPL-V, LISP) allowing the construction of formal theories of cognitive processes.

Returning to our discussion of the ecological validity of method, an example will illustrate how such a concern can impede progress in resolving questions of legitimate concern. A classic unresolved controversy concerns the affect–memory relationship. The following question was posed by several clinical and cognitive psychologists: What is the influence of the affective nature of information on memory for that information? Early research on this topic (see Rapaport, 1942/1971) sometimes showed an affective asymmetry effect (better memory for affectively pleasant than unpleasant information) and at other times showed an affective intensity effect (better memory for both pleasant and unpleasant information than for neutral information). In retrospect, it appeared that the experimenters who typically found the asymmetry effect used "real-life" memories (in particular, memories from Christmas vacation), whereas the ones that found an intensity effect had tested for memory of laboratory constructed pleasant and unpleasant information.

In the 1930s and 1940s, the ecologically valid procedure was popular partly because of the "realness" of everyday memories and also because results supported the popular psychodynamic construct of repression. It appeared not to matter much that results of research using real-life memories were not free of interpretive confounds. The most obvious confound was that pleasant and unpleasant memories may be linked to extraneous variables (such as greater spontaneous rehearsal of real-life pleasant memories) that are responsible for the observed difference rather than the affective nature of the memories per se.

Figure 1
A Two-by-Two Array of Approaches to Science



High Ecological Validity of Method but Low Generalizability of Results: Traffic Accidents, Bicycle Trips, and Oral Examinations

The procedures employed by everyday memory researchers are similar to those used in early investigations of affect and memory and, we think, are equally prone to the hazards of employing real-world settings. In our discussion, we refer to studies that appeared in the recent two-volume work, *Practical Aspects of Memory* (Gruneberg, Morris, & Sykes, 1988). These studies used real-world settings on the assumption that the results from such applications would provide information not otherwise obtainable in the laboratory.

Case 1. In a study designed to explore the relationship between stereotypes and memory of real traffic accidents, Diges (1988) interviewed people who were either

directly involved in an accident or had witnessed it. Subjects were asked to report their version of the accident, followed by a questionnaire of "driver stereotypes" (i.e., if subjects believed that women were worse drivers than men and so on). The narratives were coded for recall of (a) pre-accident or contextual ideas, (b) accident dynamics, (c) recall of people involved, and (d) recall of post-accident ideas, explanations, justifications. Correlational analyses indicated most notably that the strength of the stereotype is related to the type of recall; that is, the more extreme the belief about the driver's age or sex, the fewer the references to details about the physical aspects of the accident and the greater the references to people.

Understanding how stereotypes influence perception and memory is an important topic for investigation, and several well-conceived, well-designed laboratory studies have demonstrated that there is a link (Bruner, Busiek, & Minturn, 1952; Hastorf & Cantril, 1954; Howard & Rothbart, 1980). However, we question what can be learned about this process by (a) interviewing people whose stereotypes prior to the accident are unknown (in fact, it is quite possible that the stereotype measure obtained after the accident was influenced by the accident rather than determining the encoding of the original event, and the correlational nature of the data does not help); (b) examining the second recall, which could be contaminated by the first recall (the police report); (c) assuming no control over the original encoding of the event (in making claims about memory biases, how can it be assumed that drivers and witnesses have "seen" the same event?); (d) having only weak control over the time delay between the original event and the test; or (e) using an N that is too small (because only 13 of 500 real-world cases permitted analyses) to compare meaningfully driver and witness memory about the same event. Quite simply, the test is weak, the final results are uninterpretable, and most important, the procedure obliterates the central purpose of such a study, namely, generalizability to the real world.

Our everyday memory friends would argue that what we see as problems with the method of analyzing naturalistic memories constitute the essence of the real memories we must eventually address. Our response is that the multiplicity of uncontrolled factors in naturalistic contexts actually prohibits generalizability to other situations with different parameters. The implication that tests in the real world permit greater generalizability is false once the immense variability from one real-world situation to another is recognized.

Case 2. A study by Bruce and Read (1988) provides an example of other pitfalls of the everyday memory approach. To study memory for the frequency of naturally occurring events, Bruce (the investigator) and his wife (MVP, the subject) used events from a 47-day bicycle tour through Scandinavia. The experimenter kept records of arbitrary events, such as the number of postcards MVP mailed, traveller's checks cashed, and so on. Approximately three months later, MVP was given a memory test in which she was asked to recall the number of times each of 27 events had occurred.

MVP's responses were compared to the mean of a "control" group consisting of individuals who had not been on the trip and therefore simply guessed at how often each event was likely to have occurred. Not surprisingly, MVP showed better memory for the frequency of events on her trip than individuals who had never taken the trip. In light of the criterion of counterintuition raised by Neisser (1978), we wonder if a layperson could have correctly guessed the direction of this finding, and if so, whether this should be cause for embarrassment. Although we have not claimed that counterintuition be regarded as a criterion of good science, we do question what this finding tells us about the stated purpose of the research—understanding the how and why of memory for frequency. What it does indicate is that individuals who experienced a set of events were better at estimating the frequency of those events later than individuals who did not experience them at all.

Further, MVP was asked to indicate the process through which she arrived at each frequency estimate (i.e., I remembered the specific instances, I just made a general estimate, etc.). The data indicated that she "tended to remember and tally the specific instances of low frequency events ... [and] relied more on general impressions when judging the number of occurrences of higher frequency events" (p. 318). On the basis of these self-report ratings, the authors concluded that frequency estimates for naturally occurring events rely on the specific instances of the event, the general impression of the event, or on some combination of these two kinds of memory. Now, do everyday memory researchers really think that frequency memory in the laboratory would differ from memory for frequency information on a bike trip in Scandinavia? If, for instance, subjects were given a series of tasks containing differential frequency information to perform in the laboratory and were later tested for their memory for frequency, how would the results differ?

We suspect that had an experimental analog of this field investigation been conducted, very similar results would emerge: (a) Subjects who performed the task themselves would show higher correlations between judged and actual frequencies than subjects who were in another experiment; (b) correlations between actual and judged frequency estimates would be higher for low frequency items than high frequency items; and (c) subjects would probably report that their memory search for low frequency items involved counting each occurrence, whereas their memory for high frequency items was based on a general impression. Supposedly, the excuse for abandoning the precision of the laboratory is that there are emergent principles that one finds in real life that escape the laboratory. However, what information does such a study provide about memory for frequency that cannot be obtained in a simple laboratory experiment with better control over the tasks, a finer gradation of frequencies, incidental learning tasks, and even, perhaps, more meaningful control groups?

Retrospective reports of memory. A problem that is common to several studies of memory from naturalistic

settings, such as the two mentioned here, or to studies of memory for oral examinations from several years ago (Strube & Neubauer, 1988) is their reliance on retrospective verbal reports. Even scientists who defend the use of online verbal report data to understand complex cognitive processes such as problem solving (Ericsson & Simon, 1980) agree that retrospective reports of memory cannot be trusted. What can be the justification for accepting MVP's claim, for instance, that she searched her memory for specific occurrences for low frequency events but made a general impression of high frequency events? How is this different from asking 200 undergraduate students to tell us how they think people remember low and high frequency events? (The large *N* would at least provide us with a more substantial pool of hypotheses about memory for frequency information). Not only do everyday memory researchers often use naturally occurring contexts as their own justification, their procedures reveal the fantasy that opinions about memory expressed in real-life contexts bear an especially close resemblance to the actual workings of memory.

However, many of the everyday memory studies involve measures of the actual occurrence of events to which subjective reports could be compared. For instance, Bruce and Read (1988) reported keeping track of the actual occurrence of events, such as mailing postcards, and Strube and Neubauer (1988) had minutes of the actual oral exam available for comparison with subjects' memory of the event. Although having such information is better than not having it at all (as with analyses of dreams or childhood memories), there are problems with the method. The most obvious one is that it is impossible to separate autobiographical memory from generic information or scripts about the event. For example, in Strube and Neubauer's (1988) study of memory for oral examinations, it is possible that memory for when the oral exam occurred could be as much a function of actually remembering the time of day as it could be the knowledge that oral examinations in a certain institution were usually scheduled in the morning.

We are not surprised that field investigations of everyday memory, such as these two, succeed occasionally in turning up instances of well-known principles of memory. After all, we retain our faith that laboratory abstractions are controlled by the same laws as mundane phenomena, in psychology just as in chemistry. However, we have not been able to see any new principles of memory emerging from the everyday memory studies. Again and again, what seem at first like new, dramatic, emergent principles turn out to be everyday manifestations of laboratory wisdom. For example, early reports of flashbulb memories (Brown & Kulik, 1977) prove, on close inspection, to present nothing unexpected to conventional laboratory work on memory (McCloskey, Wible, & Cohen, 1988). Similarly, research on the role of the self in memory has shown that superior memory for self-referent information can be explained in terms of ordinary principles of memory (Greenwald & Banaji, 1989; Klein & Kihlstrom, 1987).

Principles of memory that cannot be discovered using the scientific method may indeed exist, and emergent principles may someday be discovered using everyday memory methods. One such argument, presented by Erdelyi and Goldberg (1979), was that the lack of experimental confirmation for the phenomenon of repression could not be a criterion for rejecting the idea of motivated forgetting ("existence arguments can hardly be settled on the basis of laboratory failures in creating the phenomenon—whether we are concerned with the existence of Mt. Everest, the rings of Saturn, or, for that matter, the white rat," p. 359–360). However, as Erdelyi's more recent writing indicates, Ebbinghaus may have provided the first experimental evidence for repression, which indicates that the mechanisms by which "repression" is accomplished are the same as those employed in everyday types of forgetting, such as lack of rehearsal (Erdelyi, in press). Thus, although we must reserve the possibility that everyday memory research may yield emergent principles of memory that have not and cannot be discovered in the laboratory, that supposition is, at present, a pure matter of faith.

Low Ecological Validity of Method but High Generalizability of Results

Case 1. Perhaps we can reinforce our case with examples of research findings we believe to be low in the ecological validity of method, although high in external validity as defined by the generality of conclusions. Landauer and Bjork (1978) reported such a finding. They had people memorize briefly presented paired associate items on a CRT and later tested them after measured delays. Surely Ebbinghaus himself would have been gratified by this methodology (though perhaps disapproving of the stimuli). The main manipulation was the schedule of presenting items that were exposed more than once. Departing from the distribution-of-practice literature, these authors contrived some conditions with wide spacing, some with narrow, and two new conditions with irregular spacing, one with increasing lags and the other with decreasing lags. Their main result was that increasing lags had an impressive beneficial effect on subsequent recall, comparable in size to some of the classic mnemonic techniques. As an empirical rule, the generalization seems to be that a repetition will help most if the material had been in storage long enough as to be just on the verge of being forgotten. Because repetition is known to reduce the slope of the forgetting curve, this means that successive repetitions should be scheduled with expanding lags.

Leaving aside the issue of whether this finding is painfully obvious to the undergraduate student (which it is not), we note that Landauer and Bjork could never have stumbled on this finding without using tightly controlled laboratory methods. The 50-trial sessions in which their subjects participated, sitting before the CRT, memorizing artificial first and last names, are no more ecologically representative than experiments in a bubble chamber or gas chromatography. Yet, Landauer and Bjork, as they commented, have turned in a finding that

can readily be applied to one's daily life. The technique is easily explained and can be exploited whenever one can choose the interval after which some piece of information can be rehearsed or self-tested.

Case 2. The tension between ecologically valid methods and externally generalizable principles has often recently taken the form of impatience to know what practical application justifies every popular laboratory protocol. We think this urgency denies the essential difference between science and technology. Returning momentarily to our chemical metaphor, the early alchemists had literally no idea in what directions their investigations were leading. Yet, in the course of pursuing completely the wrong questions, they stumbled on methods and ideas of great importance to the later evolution of chemistry. It may turn out to be so with the ideas we have about how our thinking about memory may be applied in the future.

A charming example of the natural evolution of science and technology came from the research program of Baddeley and his associates on working memory (Baddeley, Thomson, & Buchanan, 1975). These authors reported an elegant series of laboratory experiments showing that speaking rate affected the size of memory span for words. The longer it takes to pronounce words, the fewer of them can be remembered for immediate ordered recall, even controlling for the number of letters, syllables, and phonemes. The precision of this finding represents the best tradition in experimental control, but at the expense of an extraordinarily contrived testing situation. In a brilliant example of making derivations for everyday life, Ellis and Hennelley (1980) reasoned that because Welsh digit names are systematically longer to say than English ones, Welsh schoolchildren ought to score selectively lower than their English counterparts on digit-span indicators of IQ, but not on other indicators. They confirmed this conjecture by careful field investigations. How many other principles of memory, one wonders, are lying around in our textbooks without the benefit of insights like this to establish their everyday relevance?

For example, nowadays there is a fashion to wonder what sensory memory (iconic, echoic) is "good for" (Haber, 1983; Neisser, 1983). What everyday experience justifies the use of tachistoscopic flashes, for example? Is it wrong-headed to devote our experimental attention to methods and principles that have no obvious use in our daily lives? We think not. The justification for these efforts is to understand more adequately how vision and audition work. This understanding cannot help but further our understanding of mental processes connected with them and lead to by-products of application that are as yet unsuspected. Iconic memory, in particular, is a bad example for these critics to use because a direct outgrowth of Sperling's (1960) technique by Reicher (1969), when he discovered the word/letter effect, was responsible for opening up the whole continuing research effort exploring the microstructure of reading (e.g., McClelland & Rumelhart, 1981). Even the most cynical advocates of technology before, and in preference to, science would not deny the external validity of understanding how people read.

Strangely, Baddeley (1988) himself has been among those publicly frustrated by cases in which the immediate gratification of technological applications is not obvious for our young field. We say strangely because the Cambridge Applied Psychology Research Unit has long been among the most sensitive groups to the need for uncompromising science in the service of external generalizability. We are reminded of another early chemist who could not see the usefulness of oxygen when Lavoisier first demonstrated it in 1786. "But what the hell is it for?" he asked rhetorically, upon which many investigators halted their research while worriedly searching for the justifications of their activity. In the study of memory, we have few candidates that rival oxygen in importance, but if we look at the everyday applications of some of our principles, we are not doing too badly.

Every science goes through periods of self-analysis, and the history of any science will document the various compromises each has chosen at different times in its life cycle. We are not hard-science imperialists who believe that the only model for the study of memory is that of physics, or our rather elaborate example of chemistry. We simply believe that in a headlong rush to answer Senator Proxmire and his kind, we do not abandon the scientific method.

A Reminder of the Value of External Invalidity

Concerns about the external validity of research on memory arise in spite of several examples that point to the futility of such an exercise. A few years ago, Mook (1983) argued emphatically in defense of external invalidity, pointing out that several good reasons exist to do experiments besides predicting real-world behavior. His example of the artificiality of the laboratory setting in Harry Harlow's research is an appropriate case in point:

> And what of the representativeness of the setting? Real monkeys do not live within walls. They do not encounter mother figures made of wire mesh, with rubber nipples; nor is the advent of a terry-cloth cylinder, warmed by a light bulb, a part of their natural life-style. What can this contrived situation possibly tell us about how monkeys with natural upbringing would behave in a natural setting? ... Harlow did not conclude, "Wild monkeys in the jungle probably would choose terry-cloth over wire mothers, too, if offered the choice" ... What Harlow did conclude was that the hunger-reduction interpretation of mother love would not work. If anything about this experiment has external validity, it is this theoretical point, not the findings themselves (p. 381).

Similarly, Ebbinghaus did not intend to prepare us if, in our local grocery store (a "natural environment"), we were suddenly confronted with a memory drum presenting us with nonsense syllables at a fixed pace. He wished to inform us of the nature of forgetting and the association of ideas.

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.

One example of such a generality test, based on a prior laboratory finding was reported by Wanner (1968). He was interested in the experimental demonstration by Sachs (1967) that people soon forget surface features of text but remember the gist quite well. He tested people unexpectedly on a sentence that had been in what they assumed were the experimental instructions. This single test was the objective of the experiment; no experimental protocol even occurred after the test. The results not only confirmed Sachs's (1967) generalization, they showed the effect even more starkly. We suspect that Wanner could not, or would not, have achieved this important finding without the benefit of Sach's laboratory study in advance.

Complexity of Phenomena: The Example of Social Psychology

The nature of forgetting is a particularly troublesome question for everyday memory researchers. For instance, Neisser (1978) commented that just as it is no longer meaningful to ask questions about the causes of crime because there are so many different and complex types of crime, "forgetting is an equally incoherent notion" (p. 10). In response, we have another fundamental disagreement with believers in everyday memory methods. Our view is that the more complex a phenomenon, the greater the need to study it under controlled conditions, and the less it ought to be studied in its natural complexity. To borrow examples from social psychology, about which concerns of external validity were debated some years ago, it is because the nature of obedience to authority in Nazi Germany was known to be complex that Milgram chose to demonstrate its vicissitudes in the laboratory (Milgram, 1963). It is because the nature of bystander nonintervention in the Kitty Genovese case was complex that laboratory experiments with the power of systematic manipulations of independent variables were performed to understand it (Latane & Darley, 1970). Several of these studies incorporated laboratory manipulations that mimicked the real world and concurrently maintained the control necessary to infer a causal relationship.

In the study of memory, laboratory stimuli that approximate the real world have been frequently used. Probably the most striking examples are studies of eyewitness memory (Loftus, 1979; McCloskey & Zaragoza, 1985) in which a little red Datsun, a can of Pepsi, and hammers and wrenches are appropriate stimuli to study memory for objects of the kind an eyewitness may easily be called on to testify. However, it was because these researchers followed scientific protocol that the results have generalizability. In our view, the complexity of a phenomenon is a compelling reason to seek, not abandon, the laboratory.

The above references to social psychological experiments are reminders that the everyday memory movement's disenchantment with the laboratory is not unprecedented. An example is what came to be known in the early 1970s as the "crisis in social psychology." The crisis referred to the sentiment among many social psychologists, as well as the informed public, that social psychology had renounced its mission by being an experimental science. Who cares about the college sophomore observed though a one-way mirror filling out a 7-point scale? Social psychology, they argued in a curiously familiar voice, must be concerned with real events and real people if it is to comment on the nature of social behavior. For several years the debate continued in social psychology, but if current research procedures are any indication, the controversy was settled in favor of hypothesis derivation from theory and hypothesis testing in the laboratory.

Conclusion

In summary, we students of memory are just as interested as anybody else in why we forget where we left the car in the morning or in who was sitting across the table at yesterday's meeting. Precisely for this reason we are driven to laboratory experimentation and away from naturalistic observation. If the former method has been disappointing to some after about 100 years, so should the latter approach be disappointing after about 2,000. Above all, the superficial glitter of everyday methods should not be allowed to replace the quest for truly generalizable principles.

REFERENCES

Baddeley, A. D. (1988). But what the hell is it for? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol 1). New York: Wiley.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99.

Bruce, D. (1985). The how and why of ecological memory. *Journal of Experimental Psychology: General, 114,* 78–90.

Bruce, D., & Read, J. D. (1988). The how and why of memory for frequency. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol 1, 317–322). New York: Wiley.

Bruner, J. S., Busiek, R. D., & Minturn, A. (1952). Assimilation in the immediate reproduction of visually perceived figures. *Journal of Experimental Psychology, 44,* 151–155.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

Conrad, R. (1972). Short-term memory in the deaf: A test for speech coding. *British Journal of Psychology, 63,* 173–180.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12,* 599–607.

Diges, M. (1988). Stereotypes and memory for real traffic accidents. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol 1, pp. 59–65). New York: Wiley.

Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58,* No. 5, (Whole No. 270).

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

Ellis, N. C., & Hennelley, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71,* 43–52.

Erdelyi, M. H. (in press). Repression, reconstruction, and defense: History and integration of the psychoanalytic and experimental frameworks. In J. Singer (Ed.), *Repression: Defense mechanism and cognitive style.* Chicago: University of Chicago Press.

Erdelyi, M. H. & Goldberg, B. (1979). Let's not sweep repression under the rug: Toward a cognitive psychology of repression. In J. F. Kihlstrom & F. J. Evans (Eds.), *Functional disorders of memory* (pp. 355–402). Hillsdale, NJ: Erlbaum.

Ericsson, A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 215–251.

Greenwald, A. G., & Banaji, M. R. (1989). An experimental analog of the self as a memory system: Powerful, but ordinary. *Journal of Personality and Social Psychology, 57,* 41–54.

Gruneberg, M. M., Morris, P. E., & Sykes, R. N. (1978). *Practical aspects of memory.* London: Academic Press.

Gruneberg, M. M., Morris, P. E., & Sykes, R. N. (1988). *Practical aspects of memory: Current research and issues.* (Vols. 1 & 2). New York: Wiley.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *The Behavioral and Brain Sciences, 6,* 1–11.

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

Howard, J. W., & Rothbart, M. (1980). Social categorization and memory for in-group and out-group behavior. *Journal of Personality and Social Psychology, 38,* 301–310.

Kerr, N. H. (1983). The role of vision in "visual imagery" experiments: Evidence from the congenitally blind. *Journal of Experimental Psychology: General, 112,* 265–277.

Klatzky, R. L. (1984). *Memory and awareness: An information processing perspective.* New York: Freeman.

Klein, S., & Kihlstrom, J. (1987). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General, 115,* 26–38.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press.

Latane, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

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

McClelland, J. L., & Rumelhart, D. L. (1981). An interactive model of context effects in letter perception. Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248,* 122–130.

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,* 1–14.

McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General, 117,* 171–181.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371–378.

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). London: Academic Press.

Neisser, U. (1983). The rise and fall of the sensory register. *The Brain and Behavioral Sciences, 6,* 1–54.

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.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Rapaport, D. (1971). *Emotions and memory.* New York: International Universities Press. (Original work published 1942)

Reicher, G. M. (1969). Perceptual recognition as a function of the meaningfulness of the stimulus material. *Journal of Experimental Psychology, 81,* 275–280.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151–160.

Rundus, D. (1977). Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 665–681.

Sachs, J. D. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2,* 437–442.

Sherry, D. F., & Schachter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94,* 439–454.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74* (Whole No. 498).

Strube, G. & Neubauer, S. (1988). Remember that exam? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol 1, pp. 247–252). New York: Wiley.

Underwood, B. J. (1964). Degree of learning and the measurement of forgetting. *Journal of Verbal Learning and Verbal Behavior, 3,* 112–129.

Wanner, E. (1968). *On remembering, forgetting, and understanding sentences.* The Hague: Mouton.