

# Memory metaphors and the real-life/ laboratory controversy: Correspondence versus storehouse conceptions of memory

**Asher Koriat**

Department of Psychology, University of Haifa, Haifa, Israel  
Electronic mail: akoriat@psy.haifa.ac.il

**Morris Goldsmith**

Department of Psychology, University of Haifa, Haifa, Israel  
Electronic mail: mgold@psy.haifa.ac.il

**Abstract:** The study of memory is witnessing a spirited clash between proponents of traditional laboratory research and those advocating a more naturalistic approach to the study of "real-life" or "everyday" memory. The debate has generally centered on the "what" (content), "where" (context), and "how" (methods) of memory research. In this target article, we argue that the controversy discloses a further, more fundamental breach between two underlying memory metaphors, each having distinct implications for memory theory and assessment: Whereas traditional memory research has been dominated by the *storehouse* metaphor, leading to a focus on the *number* of items remaining in store and accessible to memory, the recent wave of everyday memory research has shifted toward a *correspondence* metaphor, focusing on the *accuracy* of memory in representing past events. The correspondence metaphor calls for a research approach that differs from the traditional one in important respects: in emphasizing the intentional-representational function of memory, in addressing the wholistic and graded aspects of memory correspondence, in taking an output-bound assessment perspective, and in allowing more room for the operation of subject-controlled metamemory processes and motivational factors. This analysis can help tie together some of the what, where, and how aspects of the "real-life/laboratory" controversy. More important, however, by explicating the unique metatheoretical foundation of the accuracy-oriented approach to memory we aim to promote a more effective exploitation of the correspondence metaphor in *both* naturalistic and laboratory research contexts.

**Keywords:** accuracy; correspondence; everyday memory; intentionality; memory; metamemory; monitoring; recall; recognition; response criterion; signal detection; storehouse:

## 1. The challenge of everyday memory

Much of the traditional laboratory research on memory conducted in the past century has followed Ebbinghaus (1895) in using tightly controlled experiments that facilitate the quantification of memory (see Baddeley 1990; Schacter 1989). This tradition has been strongly criticized in the past two decades, however, most notably by Neisser (1978), who provocatively dismissed the laboratory research of the past 100 years as largely worthless for answering "the important questions about memory" and called for a shift to the "realistic" study of memory. Since Neisser's call, there has been a growing number of studies on such varied topics as autobiographical memory, eyewitness testimony, prospective memory, "flashbulb" memory, memory for action, memory for faces, memory for places, and so forth (see, e.g., Cohen 1989; Davies & Logie 1993; Gruneberg et al. 1988; Harris & Morris 1984; Neisser & Fivush, 1994; Neisser & Winograd 1988; Rubin 1986; Winograd & Neisser 1992). This new wave of real-life or "everyday" memory research has resulted in a proliferation of research methods

that are quite removed from those traditionally used in the laboratory.

The rift between proponents of naturalistic and laboratory memory research, as well as efforts at reconciliation, may be seen in the lively debate (to which *American Psychologist* devoted its January 1991 issue) sparked by Banaji and Crowder's (1989) paper. It is apparent from the commentaries that "everyday memory" is an ill-defined category (Klatzky 1991) and that the dimensions of the controversy are not simple to specify. In general, the battles appear to be raging on three distinct fronts: *what* memory phenomena should be studied, *how* they should be studied, and *where*.

For some researchers the major issue seems to involve the content ("what") of memory research. This is reflected, for example, in the title of Neisser's (1978) leading paper, "Memory: What are the important questions?" Thus everyday memory research has been characterized by its attempt to understand "the sorts of things people do every day" (Neisser 1991, p. 35), by its choice of topics having "obvious relevance to daily life" (Klatzky 1991, p. 43), and in particu-

lar by its concern with the practical applications of memory research (e.g., Gruneberg & Morris 1992). This contrasts with the alleged irrelevance of traditional memory research, which has "chiefly focused on explicit recognition or recall of isolated items from lists" (Neisser 1991, p. 35; but see Roediger 1991).

Other discussions have treated the controversy as being over the proper research policy (the "how" question), that is, about "the most valuable ways of gaining knowledge and understanding about memory" (Loftus 1991, p. 16; see also Banaji & Crowder 1989; Tulving 1991). Proponents of the naturalistic study of memory have questioned the ecological validity of much laboratory experimentation (e.g., Aanstoos 1991), whereas laboratory proponents have stressed the importance of experimental control and the generalizability of results. Banaji and Crowder (1989), for instance, argue that because naturalistic research methods often lack experimental control, the "ecological validity of the methods as such is unimportant and can even work against generalizability" (p. 1187; see also Morton 1991; Roediger 1991). In general, naturalistic memory researchers acknowledge the desirability of controlled experimentation, but claim that a strict adherence to this methodology would leave out many interesting memory phenomena (Conway 1991; 1993; Gruneberg & Morris 1992).

Finally, still other researchers have underscored the "where" as being a fundamental, inseparable aspect of memory phenomena. For example, Neisser (1988a) has stressed the affinity between the ecological approach to the study of everyday memory and the ethological approach to studying animal behavior, both of which focus on organism-environment interactions (see also Ceci & Bronfenbrenner 1991). He therefore emphasizes the social-functional context of remembering, stating that "the theory we require will have to deal with persons, motives, and social situations. . . . Most of all, it will have to deal with functional issues" (1988b, p. 553; see also Baddeley 1988; Barclay 1993; Bruce 1985; 1989; 1991; Fivush 1988; 1993; Neisser 1978; 1991; Winograd 1988). The implication is that studying the same phenomena in the laboratory and in natural settings may lead to very different conclusions. Indeed, Gruneberg, Morris, and Sykes point to findings (Morris et al. 1985) in which "the real-life nature of the experience made a considerable difference to memory processing" (Gruneberg et al. 1991, p. 74; see also Aanstoos 1991; Bahrick 1992; Baker-Ward et al. 1993; Ceci & Bronfenbrenner 1985; 1991; Conway 1991; 1993).

It is important to note, however, that although the three dimensions - the what, how, and where dimensions - are correlated in the reality of memory research, they are not logically interdependent. For instance, many everyday memory topics can be studied in the laboratory (Neisser 1991; Roediger 1991), and memory research in naturalistic settings may be amenable to strict experimental control (Conway 1991; e.g., Ceci & Bronfenbrenner 1985; Koriat et al. 1976; Loftus 1979a). We therefore sought a further dimension of the controversy that might lie concealed behind the commonly debated issues.

We propose that the real-life/laboratory controversy harbors what appears to be a more fundamental breach - a difference in the very *metaphor* of memory implicitly espoused by each camp (see also Koriat & Goldsmith 1994; in press a). These metaphors, the *storehouse* and *correspondence* metaphors, embody two essentially different

ways of thinking about memory and how memory should be evaluated. The storehouse metaphor, which likens memory to a *depository* of input elements, implies an evaluation of the *number* of items remaining in store. In contrast, the correspondence metaphor, which treats memory as a *perception* or *description* of the past, implies an evaluation in terms of the *accuracy* or *faithfulness* of that description.

In this target article, we delineate the two contrasting metaphors and examine their respective quantity-oriented and accuracy-oriented approaches to the evaluation of memory. We believe that this analysis can help tie together some of the various aspects of the everyday-laboratory controversy. Furthermore, we contend that the distinction between the two metaphors, with their ensuing approaches to memory, is a crucial distinction in its own right, with serious implications that span the two camps. Thus, our primary aim is to explicate the unique metatheoretical foundation of the accuracy-oriented approach to memory, as opposed to the traditional, quantity-oriented approach, in order to promote a more effective exploitation of the correspondence metaphor in both naturalistic and laboratory research contexts.

The structure of the target article is as follows: In section 2, we delineate and distinguish the storehouse and correspondence metaphors and show how this distinction can explain some of the friction between proponents of everyday and traditional laboratory research. In subsequent sections we use the everyday-laboratory controversy as a backdrop, focusing primarily on the correspondence metaphor and its potential as a productive conception for memory research. In section 3 we consider how a correspondence view of memory seems to be emerging in current memory theorizing. In section 4, we explicate the unique logic of the correspondence metaphor for the evaluation of memory and outline several possible approaches to correspondence-oriented memory assessment. In section 5, we illustrate the utility of the correspondence-storehouse distinction by reviewing recent experimental work that addresses some of the troubling issues that arise in attempting to reconcile accuracy-oriented, naturalistic results with traditional, quantity-oriented laboratory findings. In particular, the correspondence metaphor is shown to call for a more serious consideration of the active role of the subject in controlling the faithfulness of his or her memory report. Finally, in section 6, we return to the everyday—laboratory controversy and outline a scheme for capturing the interrelationships between conceptual metaphors, on the one hand, and the content (what), context (where), and methods (how) of memory research, on the other. In light of our analysis of the role of metaphors in memory research, we address the issue of whether the differences between the two approaches to memory may ultimately be reconciled.

## 2. Two competing metaphors of memory

The study of memory is replete with metaphors for conceptualizing different aspects of memory and remembering (see Kolers & Roediger 1984; Malcolm 1977; Marshall & Fryer 1978; Roediger 1980; see also Gentner & Grudin 1985). Roediger (1980) compiled a "fairly complete, but certainly not exhaustive" list of 36 memory metaphors used by psychologists and philosophers from Plato until modern times - sometimes in jest, but more often quite seriously.

On the lighter side, Hintzman (1974) has compared memory ironically to a cow's stomach, and Landauer (1975) has used the even less flattering analogy of a garbage can. Although some students of memory have expressed reservations about the use of metaphors to conceptualize memory (e.g., Ebbinghaus 1895; Roediger 1980; Tulving 1979), there is no question that such metaphors have exerted a considerable influence on memory research and theory.

We concentrate here on the contrast between two general types of memory metaphors: the storehouse metaphor, which has played a dominant role in guiding traditional laboratory memory research, and the correspondence metaphor, which seems to be gaining impetus in the new wave of everyday memory research. With regard to the former metaphor, Roediger (1980) observed that "the conception of the mind as a mental space in which memories are stored and then retrieved by a search process has served as a general and powerful explanation of the phenomena of human memory. There is currently no other general conception of the mind or memory that rivals this view" (p. 238). We begin, then, with an exposition of this metaphor, focusing on its implications for memory assessment. We have chosen to present a rather strict version of the storehouse conception, providing a contrasting background for the introduction of the correspondence metaphor (in the next section). Although perhaps no investigator today would endorse such an extreme version, it is nonetheless important to confront its implicit logic, which still pervades much contemporary research and thinking about memory. In this regard, we subscribe to the rationale offered by Jussim (1991) for his critical analysis of the strong constructivist perspective in social perception: "Regardless of whether anyone actually believes in the [strict storehouse metaphor], clearly many choose research topics, write, and interpret research as if they believed it" (p. 55, brackets substituted for original text).

### 2.1. The storehouse metaphor and its implications

Despite the skepticism expressed by Ebbinghaus himself about the utility of memory metaphors, much of the Ebbinghaus tradition of laboratory memory research has been guided by a metaphor of memory as a storehouse of discrete, elementary "units." The origin of this metaphor may be traced back as far as Plato (see Herrmann & Chaffin 1988; Marshall & Fryer 1978), but its more modern development may be seen in British empiricist philosophies and their associative-atomistic conception of the mind as a store of elementary "ideas" and "associations" (see Mandler & Mandler 1964; O'Neill 1968). Thus, according to Locke (1690), memory "is as it were the storehouse of our ideas. . . . a repository to lay up those ideas" (Book 2, Ch. 10). In this conception, a multitude of stimuli is assumed to impinge upon the senses, and discrete impressions of these stimuli are retained as memory units for later retrieval. As a result of decay or interference some of the units may become lost, weakened, or otherwise inaccessible.

With regard to memory assessment, the storehouse metaphor has legitimized the use of discrete, elementary stimuli as experimental input, allowing for the *quantification* of memory (Schacter 1989). This approach permeates a vast number of studies carried out in the past century using a variety of experimental paradigms and memory measures (see, e.g., Crowder 1976; Gregg 1986; Murdock 1974).

Indeed, the prototypical list-learning paradigm, the workhorse of the laboratory tradition, essentially simulates the course of events assumed to take place when the input elements are first "deposited" in the memory store and later "recollected" or "retrieved." The stimuli, typically referred to as "items," consist of nonsense syllables, words, and so on, whose salient characteristic is their *countability*—they allow measures of memory effectiveness based on the *number* of recovered elements. Implicit in this approach is the conception of forgetting as *information loss*, either the loss of the elementary units themselves (item information) or the loss of the associative links between them (associative information; see Murdock 1974). Thus the most natural measure of memory is simply how many of the units of information originally presented can be recovered on a given memory test. In fact, in the great majority of laboratory studies on free recall, incorrect responses (i.e., commission errors) are simply ignored. Also, it makes no difference, for instance, whether HAT was remembered and GUN was forgotten, or vice versa: all elements retrieved from the memory store are equivalent, that is, interchangeable as far as the total memory score is concerned. In sum, what matters is not *what* is remembered, but rather *how much*.

The storehouse metaphor, with its associated *quantity-oriented* approach to memory, has had a pervasive impact on the mainstream of traditional memory research. It has directed researchers' thinking toward such aspects of memory as storage capacity, the internal architecture of the store, the transfer of units from one department to another, competition between units, and, of course, information loss. It has also shaped both the experimental paradigms used to study memory (e.g., list learning, paired associates, etc.; see Puff 1982) and the types of phenomena investigated (e.g., the effects of list length, retention interval, spacing, serial order, etc.). Moreover, although the storehouse conception is perhaps most firmly rooted in the verbal learning tradition, its influence has extended to other research traditions as well. Significantly, this conception appears to have been well-suited for the more recent, information-processing framework and its associated metaphor, the digital computer (see Lachman et al. 1979; Tulving 1979). The computer metaphor adds greater sophistication in terms of internal organization, coding, processing, and transformation, yet it has served primarily to reinforce the fundamental storehouse features - the input, storage, and retrieval of discrete units of information. In fact, the "modal model" of separate memory stores (e.g., Atkinson & Shiffrin 1968) appears to represent but one example of a more modern generation of computerized storehouses (see Marshall & Fryer 1978; Roediger 1980).

It may be seen, then, that as memory research has progressed, the "pure" strain of storehouse metaphor has evolved into a variety of related species, adapted to fit the different requirements of different memory phenomena. For instance, some discussions imply a differentiated, organized storehouse made up of many departments, with items stocked under specific addresses (cf. "library" or "dictionary" metaphors; Broadbent 1971; Loftus 1977; Marshall & Fryer 1978). Others imply the stacking of items in layers, one on top of the other, so that "buried" items are more difficult to reach (Bekerian & Bowers 1983). Still others assume that some of the departments are more limited in space than others, so that items must be pushed from one to

another (Atkinson & Shiffrin 1968; Waugh & Norman 1965). Also, the "items" themselves may be more or less complex, ranging from simple features, verbal units, ideas and associations, to propositions, chunks, and templates (see Malcolm 1977).

Nevertheless, although modern treatments of memory have come a long way from the simplistic storehouse conception, many of the respected theories and experimental paradigms still bear the stamp of their storehouse ancestry in adhering to the quantity-oriented approach to memory (see Schacter 1989). Indeed, the staying power of the storehouse metaphor remains apparent not only in the models themselves, but also in the classical experimental tools that continue to dominate laboratory memory research, namely, the "memory drum" and its modern descendants. We should point out, however, that these tools and paradigms are often used with aims different from those for which they were originally intended (see sect. 4.2), and in connection with memory phenomena that would appear to call for a different type of metaphor. Not coincidentally, then, laboratory procedures that seem to break with the storehouse mold have met with relative approval from advocates of everyday memory research, who admit that they are "appreciably closer (than the old methods were) to the sorts of things people do everyday" (Neisser 1991, p. 35). In fact, laboratory proponents have had to contest the placement of entire areas of laboratory research under the banner of everyday memory, such as "research on the tip-of-the-tongue phenomenon, . . . the feeling-of-knowing experience, . . . eyewitness testimony, . . . and reality monitoring" (Roediger 1991, p. 38). Such research does not just deal with memory phenomena that occur everyday: We suggest that there is an additional, more fundamental aspect that tends to distinguish such studies from the traditional laboratory approach - the implicit metaphor of memory on which they are based.

## 2.2. The correspondence metaphor and its implications

Despite its dominance in guiding traditional laboratory research, the storehouse metaphor would seem to have limited value for the study of many everyday memory phenomena. Consider, for example, a situation in which a person on the witness stand is asked to report what she can remember about the circumstances of a crime. This situation, like many other real-life situations, motivates a different way of thinking about memory, one in which the intrinsic quality of memory is not its storage capability, but rather its ability to faithfully represent the past. Thus, the basic criterion for evaluating memory is not the quantity of items remaining in store, but rather, the *correspondence* between what the person reports and what actually happened (see Winograd 1994). Unfortunately, there appears to be no single concrete metaphor (like the storehouse) that alone can provide the essential features for such an alternative conception. Therefore, we have chosen to explicate a more abstract correspondence metaphor (Koriat & Goldsmith 1994; in press a) with the following interrelated attributes:

First, memory is conceived as being *about* some past event, to constitute a *representation* or *description* of the past episode (see, e.g., Conway 1991; 1993). Consequently, memory reports are treated as *propositions* that have a *truth value*, that is, that can be judged as right or wrong, or

as being more or less "true" to aspects of the actual event (e.g., the actual speed of the car).

Second, as just stated, the essential feature of memory is its ability to faithfully represent past events. Thus, memory is evaluated in terms of its *accuracy*, that is, its "fit" with past events, the extent to which it accords with reality (or some other criterion; see Ross, in press), rather than in terms of the number of items remaining in store. Likewise, forgetting is conceived as a *loss of correspondence* between the memory report and the actual event, that is, as a *deviation from veridicality*, rather than as just a loss of items. Thus, this conception entails a unique concern with the many different types of *qualitative memory distortions* - fabrication, confabulation, simplification, and the like (see, e.g., Alba & Hasher 1983; Bahrack et al. 1993; Bartlett 1932; Brewer & Nakamura 1984; Dawes 1966; Goldmeier 1982; Loftus 1979a; 1979b; 1982; Neisser 1981; 1988c; Riley 1962; Wells & Loftus 1984).

Third, memory correspondence is *content* laden. Unlike the storehouse metaphor, which engenders a predominant concern with *how much* is remembered, the correspondence metaphor (and virtually all real-life memory situations) entails an additional concern with the *quality* of memory (Schacter 1989), that is, with *what* is remembered (Conway 1991; 1993). In the courtroom, for instance, it might make a crucial difference whether the witness remembers that the burglar "had a gun," but forgets that he "wore a hat," rather than vice versa. Thus, *functional* considerations are intrinsic to the evaluation and study of memory correspondence (see below).

Fourth, in contrast to the evaluation of memory "storage," the evaluation of memory correspondence is inherently *output bound*: rather than begin with the input and ask how much of it is recovered in the output, one naturally begins with the output (i.e., the memory report) and examines to what extent it accords with the input. In general, accuracy is meaningful only for what a person reports (e.g., the color of a shirt, the speed of a car), not for what is omitted. Thus, while under the storehouse view subjects are held accountable primarily for what they *fail* to report, under the correspondence view subjects are accountable primarily for what they *do* report.

Finally, the correspondence conception of memory has much in common with the way we think about *perception*. In perception, interest lies in the correspondence between what we perceive and what is out there, that is, in the (output-bound) *veridicality* of our perceptions, and in the various ways in which they may deviate from reality (e.g., illusions). Likewise, under the correspondence metaphor, memory may be conceived as the *perception of the past*, and the question then becomes to what extent this perception is dependable (cf. "memory psychophysics," Algom 1992).

Indeed, many of the metaphors underlying perceptual theory would appear to imply their counterparts in correspondence views of memory. Just as perception has been viewed alternatively as a passive reflection of the external environment (Locke 1690), as an active construction of reality (e.g., Neisser 1967; Rock 1983), or as a direct "resonance" to ecological "affordances" (Gibson 1979), likewise memory may be conceived as mirroring past experience (see Brewer, in press; Malcolm 1977), as an active reconstruction of past events (e.g., Bartlett 1932; Neisser 1967), or even as a "stage-setting" attunement (Bransford et al. 1977). With regard to assessment, under the correspon-

dence metaphor — as in perception — one is generally not concerned with how much of the impinging information is remembered (perceived), but rather with the output-bound correspondence or "goodness of fit" (see sect. 4.1) between what is remembered (perceived) and what actually occurred.

Collectively, these aspects of the correspondence metaphor characterize what we have called an *accuracy-oriented* approach to memory (Koriat & Goldsmith 1994a; in press a). This approach is reflected in much of the work on everyday memory, particularly in areas such as autobiographical memory and eyewitness testimony, which disclose a pervasive preoccupation with the faithfulness and dependability of memory for past events (e.g., Barclay 1988, 1993; Barclay & Wellman 1986; Brewer 1988; in press; Deffenbacher 1988; 1991; Hilgard & Loftus 1979; Loftus 1979a; 1979b; 1982; Neisser 1981; 1988b; Neisser & Fivush, 1994; Ross 1989; in press; Ross & Buehler, 1994; Rubin 1986; Wells & Loftus 1984; Winograd, 1994; Winograd & Neisser 1992).

This preoccupation is not arbitrary. The affinity between the correspondence metaphor and everyday memory research appears to stem from the basic character of memory in everyday life, where *what* is being remembered is certainly no less important than how much, and where memory reports are naturally considered to be *about* personally experienced, past events and states (see Conway 1991; 1993). The difference between treating memory as being about something (correspondence metaphor) and treating it as the mere retrieval of something (storehouse metaphor) is so obvious that it can easily be overlooked. A recalled list of words, for instance, need not be considered as being about anything; it can simply be treated as the retrieval of the items that remain in store. Thus, as Neisser (1981) observed, traditional laboratory research has generally studied memory for "material that has no reference beyond itself" (p. 4). By contrast, stressing the *intentionality* ("aboutness") of real-life memory, Conway (1991) goes as far as to propose that the study of everyday memory may require a *different theory of mind* than one which would have us "study human memory as if it were a chemical reaction - like dough rising." He asserts that "one difference between mental and physical states is that mental states have content, whereas physical states do not. Thus, my memory of dough rising is *about* something, some representation of an event I once experienced. But actual dough rising is not *about* anything; it is simply what it is - dough rising" (p. 24, emphasis in original).

Conway's remarks reflect a view where memory does not serve merely as a depository of isolated, lifeless units, but rather affords a meaningful representation of real-life events that can be effectively utilized in future interactions. This *functional* perspective (e.g., Baddeley 1988; Barclay 1993; Bruce 1985; 1989; 1991; Fivush 1988; 1993; Neisser 1978; 1988b; 1991; Nilsson 1979; Winograd 1988) motivates a concern for the *dependability* of memory, that is, the extent to which it can be counted on to faithfully reflect past events.

Furthermore, when memory is viewed with reference to past events, it becomes clear that memory representations can deviate from reality in many different ways. Thus, the recent wave of everyday memory research has brought a renewed interest in memory errors, particularly in the qualitative changes that occur in memory for complex,

meaningful material (Bartlett 1932). This interest is inherent to the correspondence view. Indeed, because real-life experience is made up of richly structured scenes and events (see McCauley 1988, Neisser 1986; 1988c), many naturalistic errors pertain to wholistic and relational changes that cannot be readily captured in terms of the mere loss of "items" (see, e.g., Alba & Hasher 1983; Bartlett 1932; Brewer & Nakamura 1984; Dawes 1966; Neisser 1981; 1986; 1988c). Such changes also may reflect social, motivational, and functional biases that are quite foreign to the passive storehouse conception (see, e.g., Boon & Davies 1988; Neisser 1988a; 1988b; Nigro & Neisser 1983; Ross 1989; Ross & Buehler, 1994).

At the extreme, the concerns of correspondence-oriented and storehouse-oriented researchers may be so different as to seem almost unbridgeable. Consider, for example, the following quote from Neisser (1981) regarding the quality of John Dean's memory (emphasis added in order to highlight the correspondence way of thinking): Analysis of Dean's testimony does indeed reveal some instances of memory for the *gist* of what was said on a particular occasion.

Elsewhere in his testimony, however, there is surprisingly little *correspondence* between the course of a conversation and his *account* of it. Even in those cases, however, there is usually a *deeper level* at which he is right. He gave an *accurate* portrayal of the real situation, of the actual characters and commitments of the people he knew, and of the events that lay behind the conversations he was trying to remember. Psychology is unaccustomed to analyzing the *truthfulness* of memory at this level, because we usually work with laboratory material that has no *reference* beyond itself, (p. 4)

We leave it to the reader to consider how the concerns expressed by Neisser might be accommodated within the storehouse conception.

In sum, there seems to be more brewing within the real-life/laboratory controversy than just the what, how, and where issues. Clearly, the rise of the everyday memory camp does not stem from a disdain of controlled experimentation, nor is it simply a reaction against the laboratory context as such. Rather, we contend that it reflects, at least partly, an acute disillusionment with the kind of *thinking* about memory that has permeated the traditional laboratory approach. This particular way of thinking - the storehouse conception - is embodied in the established laboratory tools and paradigms used in the quantity-oriented study of memory. Of course the relationship between the metaphors and the everyday—laboratory affiliations is more a correlation than a perfect mapping: Much work on memory correspondence has been (and hopefully will be) conducted by laboratory researchers (see sect. 3), and much everyday memory research continues to submit to the alluring power of the storehouse metaphor (see sect. 6). Nonetheless, the correspondence metaphor sketched above consolidates many of the objections levelled against traditional memory research and seems to underlie the type of accuracy-oriented approach now gaining impetus in the study of everyday memory.

In the following sections, then, we focus on the emerging correspondence conception of memory and examine its potential as a viable alternative to the storehouse conception in providing a productive framework for memory research. In doing so, we have several aims: (1) to document the emergence of the correspondence metaphor in current memory theorizing; (2) to explicate the unique logic of the correspondence metaphor and to pursue its implications

for the study of memory; (3) to show how the storehouse-correspondence distinction can help clarify some troubling issues that arise when comparing laboratory (quantity-oriented) and naturalistic (accuracy-oriented) research findings; and (4) to illustrate, more generally, the way in which a conceptual metaphor can help shape both the theories and the methods of scientific research.

### 3. The correspondence metaphor in memory research and theory

In the foregoing discussion we argued that the new wave of everyday memory research discloses a correspondence-oriented approach to memory that differs fundamentally from the storehouse-oriented approach that has dominated traditional laboratory research. The correspondence view of memory, however, appears to be gaining influence in current laboratory-based research and theorizing as well. Indeed, signs of a general shift away from storehouse-guided theorizing and toward a correspondence-oriented metatheory may be discerned in a wide variety of contemporary approaches, including the reconstructive, attributional, ecological, functional, nonmediational, procedural, and connectionist approaches to memory.<sup>1</sup>

We now consider each of these approaches in turn and briefly discuss how each seems to manifest different facets of the correspondence metaphor.

The first serious proposal for a correspondence-oriented view of memory was advanced by Bartlett (1932), who wrote "remembering is not the re-excitation of innumerable fixed, lifeless and fragmentary traces. It is an imaginative reconstruction or construction" (p. 213). Bartlett conceived of remembering as an attempt to *make sense* of experience by applying cognitive structures, called "schemata" (a concept introduced by Head, 1920). These structures constitute "an active organization of past reactions, or of past experiences" operating as a "unitary mass" (Bartlett, 1932, p. 201). Bartlett's reconstructive approach to memory was given further impetus by Neisser (1967),<sup>2</sup> and today this approach clearly encompasses a substantial amount of both everyday and laboratory memory research. Indeed, many current theoretical notions, such as schemata, frames, scripts, plans, MOPs, TOPs, mental models, and story grammars (e.g., Johnson-Laird 1983; Kintsch & van Dijk 1978; Mandler 1979; Minsky 1975; Rumelhart 1975; 1980; Schank 1982; see also Rumelhart & Norman 1988) reflect the basic assumption that remembering is an active, constructive "effort after meaning" (Bartlett 1932, p. 20).

The implications of the reconstructive view have been investigated experimentally using a wide range of rich and complex stimulus materials and tasks, including memory for sentences, stories, and real-life events (for reviews, see Alba & Hasher 1983; Brewer & Nakamura 1984). In eyewitness research, for instance, Loftus (1979a; 1979b; 1982) and her colleagues have been very influential in demonstrating the many ways in which memory for witnessed events can be distorted by reconstructive inference, particularly inference based on postevent information (e.g., Loftus et al. 1978; Loftus & Palmer 1974). Also, in autobiographical memory research, the "self-schema" (Markus 1977) has been used extensively to explain both the accuracy and inaccuracy of memory constructions for personally experienced states and events (e.g., Barclay 1986; 1988;

1993; Barclay & Wellman 1986; Markus 1980; Neisser 1988b; Ross 1989; Ross & Buehler, 1994; Winograd 1994). More generally, the reconstructive view has inspired the postulation and study of a variety of selective, integrative, and interpretive processes in memory (e.g., Bower et al. 1979; Bransford & Franks 1971; Bransford & Johnson 1972; Dooling & Christiaansen 1977; Johnson et al. 1973; Morris et al. 1979; Pichert & Anderson 1977; Seifert et al. 1985; Spiro 1980; Wagenaar & Boer 1987; but see Alba & Hasher, 1983, for reservations).

On the whole, the reconstructive approach goes far beyond the storehouse conception of memory in emphasizing the active role of the rememberer in creating a meaningful and organized representation of past events and in admitting a variety of qualitative ways in which this representation can deviate from reality. This approach, then, is perhaps the clearest and most productive example of a contemporary, correspondence-oriented approach to memory.

A fundamental criticism that has been leveled against the reconstructive approach, however, is that it has not moved far enough away from the atomistic and mediational assumptions of the storehouse metaphor. For instance, Marshall and Fryer (1978) contend that "currently, Bartlett is often cited in support of the notion that 'much of what is remembered is reconstructed from stored fragments' (Fodor 1975). . . . but this is not to impugn the storehouse metaphor at all, it is merely to offer a variation on its contents" (p. 8). This criticism is perhaps somewhat overstated. Bartlett himself was emphatic in dissociating his ideas from the storehouse metaphor, stressing that "a storehouse is a place where things are put in the hope that they may be found again when they are wanted exactly as they were when first stored away. The schemata are . . . living, constantly developing, affected by every bit of incoming sensational experience of a given kind. The storehouse notion is as far removed from this as it well could be" (1932, p. 200).<sup>3</sup> However, as will be discussed later, this criticism does seem to hold for some of the specific experimental practices and evaluative procedures employed by students of memory reconstruction, among others (see sect. 4).

More far-reaching departures from the storehouse view of memory seem to be emerging, however. For example, Jacoby and his associates (e.g., Jacoby 1988; Jacoby et al. 1989; 1992; Kelley & Jacoby 1990) have promoted a constructive-attributional approach that places a special emphasis on the subjective experience of remembering. This experience is seen to result from attributions to the past that are evoked by a feeling of familiarity for a present stimulus. When the fluent processing of a present stimulus, and hence its subjective familiarity, actually does derive from previous exposure, then the attribution to the past should result in correct or veridical remembering. However, a feeling of familiarity may also derive from other sources, and when improperly attributed to the past, may give rise to confabulations and memory illusions. Thus, in investigating the genesis of memorial experiences, Jacoby and coworkers have demonstrated how false memories may be created by altering perceptual processes independently of past experience. In their view, "the conscious experience of remembering is not to be found in a memory trace. Rather, remembering is an inference based on internal and situational cues" (Kelley & Jacoby 1990, p. 49). This work, then, displays a concern with both veridical and nonveridical

remembering, and suggests possible reasons why people's current "perceptions" of the past might deviate from reality.<sup>4</sup>

Another road, which also stresses the link between perception and memory, has been taken by researchers advocating an ecological or "direct" approach to memory, along the lines proposed by Gibson (1979; see, e.g., Bransford et al. 1977; Neisser 1986; 1988a, 1988c). Neisser (1986), for instance, renounced his earlier (1967) conception of memory reconstruction for a more direct, ecological view: Events and extenditures [extended events] have nested levels of structure, and so does our experience of them. How does that correspondence come about? This question does not even arise in traditional information-processing theories, because they do not discuss the real environment at all. They simplify matters by assuming that, whatever the world may actually be like, only very molecular information about it is available to perceivers. . . . The ecological approach assumes, in contrast, that molar events are often perceived just as directly as molecular ones. The environment is equally real at different levels of analysis, and events at those different levels are typically specified by different kinds of information. . . . It is my hypothesis, that these several levels of analysis are each represented in memory, leaving more or less independent "traces" behind, (pp. 75-76)

Neisser, then, emphasizes the *structured correspondence* between perception and memory, on the one hand, and reality, on the other. He also suggests a natural starting point for investigating that correspondence: "Rather than beginning with the hypothetical models of mental functioning, ecological psychologists start with the real environment and the individual's adaptation to that environment" (Neisser 1988a, p. 153).

The concern with functional issues expressed by Neisser represents another manifestation of current disillusionment with the storehouse metaphor. In fact, the functional approach to memory tends to be associated with two somewhat different emphases (see Winograd 1988). The first, of course, is the heightened concern with memory function itself, that is, what memory is *for* (Baddeley 1988; see also Barclay 1993; Bruce 1985; 1989; 1991; Fivush 1988; 1993; Neisser 1978; 1988a; 1988b; 1991; Nilsson 1979; Sherry & Schacter 1987; Tulving 1985). This concern may range from the adaptive role of memory at an evolutionary scale (e.g., Bruce 1985; Sherry & Schacter 1987; Tulving 1985) down to the level of social or individual interests that are served (e.g., Barclay 1993; Fivush 1993; Neisser 1981; 1988b). As we shall see (in sect. 4), such functional considerations are also critical in developing the kind of memory measures that follow uniquely from the correspondence metaphor.

The second aspect associated with the functional approach is a deemphasis of concern with the mediating mechanisms of memory. For instance, Bahrlick (1987) explains, "by functional approaches, I refer to theories that attempt to establish parsimonious relations between manipulated variables and memory performance, without necessarily attempting to reach conclusions about internal processing" (pp. 389-90; also see Jacoby 1988). Proponents of a nonmediational approach to memory have gone so far as to reject entirely the existence of a mediating "substrate of memory" (Watkins 1990), particularly as it is embodied in the concept of "memory trace" (see also Bransford et al. 1977; Craik 1983; 1991; Crowder 1993; Kolers 1973; Kolers & Roediger 1984; Kolers & Smythe 1984; Lockhart & Craik

1990; Malcolm 1977). Such views imply a revolt against the storehouse metaphor, whose input-bound perspective naturally leads to a concern for the fate of the memory trace.

An additional proposal comes from those subscribing to a "proceduralistic" view of memory (e.g., Craik 1983; 1991; Crowder 1993; Kolers 1973; Kolers & Roediger 1984; Kolers & Smythe 1984; Lockhart & Craik 1990; Roediger, Weldon et al. 1989), in which memory is assumed to involve the retention of information by the same units that processed it originally. According to this view, "the memory trace should be understood, not as the result of a specialized memory-encoding process, but rather as a by-product or record of normal cognitive processes such as comprehension, categorization, or discrimination" (Lockhart & Craik 1990, p. 89). This conception is based on evidence showing "that the means of acquisition of information form part of its representation in mind, that recognition varies with the similarity of procedures in acquisition and test, and that transfer between tasks varies with the degree of correspondence of underlying procedures" (Kolers & Roediger 1984, p. 425).

Several facets of the proceduralistic view may be especially valuable in advancing theories of memory correspondence. The first is the emphasis on the context of remembering, in particular on the *congruence* between the relevant processing activities of the rememberer at the time of remembering and at the time of witnessing the original event. Although laboratory research has typically focused on the importance of contextual congruence in affecting the *quantity* of information that can be recovered (e.g., the effects of "encoding specificity" and "state-dependent learning"; see, e.g., Eich 1980; Fisher 1981; Fisher & Craik 1977; Mantyla 1986; Morris et al. 1977; Smith et al. 1978; Tulving & Thomson 1973; Watkins & Tulving 1975), such congruence may also be critical in determining the *correspondence* between remembered and actual events (see, e.g., Fisher & Geiselman 1992; Geiselman & Fisher 1989). The second important aspect is the emphasis on the *active role of the subject*, both in determining the way in which information is initially processed (elaborated, etc.) and in directing the subsequent act of remembering (e.g., the "levels of processing" approach; see Craik & Lockhart 1972; Lockhart & Craik 1990). A third consideration is the conception of memory as embodied in *global changes* to the response tendencies of the rememberer, rather than in the storage of discrete memory traces. This conception also stresses the affinity between memory and perception:

If remembering is closely akin to perceiving, then it is perhaps no more likely that memory traces exist in the absence of remembering than percepts exist in the absence of perceiving: The activity must be studied while it is occurring. Clearly something in the system must change as a result of experience, but the changes may be diffuse and widespread modifications of the whole cognitive system so that the system now interacts with aspects of the environment in a different way, rather than events being recorded specifically and discretely like events on a video recorder. (Craik 1983, p. 356)

A similar conception, of course, is intrinsic to the connectionist, parallel distributed processing (PDP) approach to modeling human memory (e.g., Hinton & Anderson 1981; McClelland & Rumelhart 1986a; 1986b; Rumelhart & McClelland 1986; for a review, see Hintzman 1990). This approach, based on a *brain* metaphor (Rumelhart 1989),<sup>5</sup> offers a promising vehicle for the development of corres-

pondence-based memory models. In PDP models, knowledge is not "stored" in any specific location, but rather is embodied in the connections between a multitude of interacting processing units. Both input and output are represented as complex patterns of activity, and the system learns by adjusting the connection weights, not by incorporating new discrete elements or representations. Thus memory is distributed across the entire system, and learning reflects changes in the response tendencies of the system as a whole.

Connectionist models treat memory, like perception, as essentially a problem in pattern recognition, based on a principle of global matching (see Hintzman 1990). Thus the accuracy of memory responses in such models is most naturally evaluated in terms of the overall correspondence between distributed patterns of activity, rather than in terms of the mere amount of information recovered (e.g., "resonance scores," Metcalfe 1990; Metcalfe & Eich 1985). This approach is also well suited for investigating memory distortions (e.g., memory "blending"; Metcalfe 1990). Indeed, both the way in which new inputs are incorporated into the existing network structure (assimilation), and the gradual modification of this structure that results from the processing itself (accommodation), make connectionism a natural choice for modelling the wholistic, schematic, and constructive aspects of memory processing (see, e.g., Grossberg 1987; Hintzman 1986; McClelland & Rumelhart 1985; 1986b; Metcalfe 1990; Rumelhart et al. 1986).<sup>6</sup> Moreover, connectionism may offer a solution to the troubling issue of intentionality or "aboutness" (Bechtel 1988; Bechtel & Abrahamsen 1991) that plagues traditional cognitive models (cf. the "symbol grounding" problem; Harnad 1990). As argued by Bechtel and Abrahamsen (1991), because connectionist representations "constitute the system's adaptation to the input, there is a clear respect in which they would be *about* objects or events in the environment if the system were connected, via sensory-motor organs, to such an environment" (p. 129, emphasis in original; see also Harnad 1990; 1992).

In sum, although quite varied in many aspects of their conception of memory, the approaches considered in this brief survey all seem to reflect a basic shift from the traditional storehouse conception toward a correspondence-oriented perspective, emphasizing the congruence between current remembering and past events, and specifying factors that may affect this congruence. Many of the approaches also embody a view of memory as the perception of the past and are specifically concerned with the veridicality of this perception. Thus, although the preoccupation with memory correspondence is most salient in everyday memory research, the emergence of the correspondence view of memory would seem to represent a broad undercurrent that is gaining momentum in the study of memory generally.

It is curious, however, that despite the growing theoretical interest in memory correspondence, many of the methodological practices applied to the study of memory correspondence still seem to pay homage to the storehouse metaphor in the way that memory is *evaluated*. In fact, as we have noted elsewhere (Koriat & Goldsmith, in press a), although many memory researchers today *talk* correspondence, they still *practice* storehouse. A couple of examples should suffice to illustrate this point. The reconstructive view of memory, for instance, might have been expected to

promote the development of memory measures that uniquely reflect the kind of global correspondence (or miscorrespondence) between current memories and past events that is assumed to ensue from reconstructive processes. Indeed, Bartlett (1932) firmly rejected the applicability of Ebbinghaus-style memory measures. Yet, many predictions of the reconstructive view are commonly tested by focusing on the quantity of discrete items of information remembered (e.g., Bransford & Johnson 1972; Brewer & Treyns 1981; Dooling & Mullet 1973; Kozminsky 1977; Morris et al. 1979; Pitched: & Anderson 1977). Also, in naturalistic accuracy-oriented research, many of the memory measures are still based on the number of stimulus items correctly recalled or recognized, for instance, the number of correct propositions about a witnessed event (e.g., Fisher et al. 1989; Ornstein et al. 1992; see also Hilgard & Loftus 1979), the number of correct mugshot or lineup identifications (e.g., Brown et al. 1977; Gorenstein & Ellsworth 1980; Lindsay & Wells 1985; see also Deffenbacher 1991), and so forth. Although valuable information can certainly be gained from such measures, they do not capture many aspects of memory that are of unique concern in the study of memory correspondence.

There appears to be a serious gap, then, between the concern with memory correspondence, on the one hand, and the standard methods by which memory is in fact evaluated, on the other. This gap may derive from the failure by students of memory to realize that the focus on the faithfulness of memory implies a different memory metaphor - and hence a different approach to memory assessment - than the traditional focus on the quantity of remembered information. Indeed, it is unfortunate that in contrast to the storehouse-guided, quantity-oriented assessment of memory which has benefitted from a great deal of methodological analysis, little systematic effort has been invested in clarifying the unique logic underlying the evaluation of memory correspondence. This lack of methodological base can lead to difficulties and confusions in comparing research findings that cut across the quantity-oriented and accuracy-oriented approaches (see sect. 5). Even more important, it is our belief that a careful examination of the unique logic underlying correspondence-oriented memory assessment can provide a first step toward a more effective exploitation of the correspondence metaphor in the practice of memory research, both in the laboratory and in naturalistic settings.

#### 4. The correspondence-oriented evaluation of memory

In this section, we examine the implications of the correspondence metaphor for the evaluation of memory and consider several ways this metaphor can actually be implemented in memory assessment. We distinguish between two general approaches: the *analytic* approach assumes that memory (or at least memory reports) can be meaningfully sliced into individual, isolated units. This assumption is shared with the storehouse conception, and therefore its adoption in the context of the correspondence metaphor allows both accuracy-based and quantity-based measures of memory to be derived and compared within a common framework. This is the approach taken in our own research, which will be sketched out in section 5.

However, it is in the context of the *wholistic* approach that the correspondence metaphor finds its most unique expression. This approach attempts to avoid the segmentation of experience into separate units and strives to reach an overall measure of correspondence for the memory output as a configured whole. We begin, then, with a discussion of this approach, which best illustrates the unique flavor of correspondence-oriented memory assessment.

#### 4.1. The wholistic evaluation of memory correspondence

The logic of the storehouse metaphor is well reflected in the so-called "forgetting functions," which convey the percentage of items recoverable at any point in time. Turning to the correspondence metaphor, are there accuracy measures that can allow us to plot analogous correspondence-based functions, conveying reductions in the *overall faithfulness* of memory? Unfortunately, the derivation of such measures is no simple task. Once it is admitted that forgetting involves more than the mere omission of items, memory assessment becomes complicated by the fact that memory reports can deviate from the original event in many different ways. Thus, memory researchers emphasizing the qualitative changes that occur in memory have generally confined themselves to the study of specific types of distortions (see, e.g., Bartlett 1932; Dawes 1966; Goldmeier 1982; Loftus 1982; Riley 1962) rather than tackle the serious problems inherent in deriving an overall "faithfulness" measure that cuts across various dimensions of miscorrespondence (but see Neisser & Harsch 1992; Neisser et al. 1991). What would the development of such a measure entail?

Consider a simple situation where several persons are exposed to the same event and singly asked to report what they remember. How should an experimenter, who wishes to quantitatively evaluate the overall correspondence of each report to the actual event, proceed?

First, the experimenter must identify the relevant aspects or dimensions of the event, their relative importance (weight), and the relative criticality of different types of distortion and error. Clearly, these decisions will depend on the given memory domain and will need to incorporate functional considerations pertaining to the reasons for remembering and the particular circumstances of the memory report (see Neisser 1988b; 1988c). Furthermore, it may be necessary to take into account the level of detail or "grain" of the reports, as the correspondence of memory reports to past events can increase dramatically when responses are more general and less detailed (Neisser 1988b; Yaniv & Foster 1990; 1995). It also might be helpful to have some theory of the various ways in which memory can go wrong, to aid in the identification and measurement of the relevant dimensions of miscorrespondence (e.g., Bartlett 1932; Dawes 1966; Goldmeier 1982).

Second, given that such decisions have been made, a *quantitative assessment model* will need to be developed in order to allow for the computation of an overall correspondence score. This model must specify the various aforementioned aspects, as well as operationally define (1) how each dimension of correspondence is to be measured (e.g., Mandler & Johnson 1977), (2) how the correspondence score should be integrated across dimensions (e.g., Neisser & Harsch 1992), and (3) how differences in the "grain size" of the reports are to be taken into account (e.g., Yaniv & Foster 1990; 1995).

A major obstacle to the development of wholistic correspondence measures, then, is that, unlike the traditional measures of memory quantity (e.g., percent correct on a free-recall task) that can be applied across a wide variety of testing situations, overall measures of memory faithfulness may need to be *theory-specific*, *domain-specific*, and *function-specific*. (A similar view has been expressed with regard to evaluating the accuracy of social perception; see Kruglanski 1989.) This, of course, may tend to limit the generalizability of results that are obtained with such measures (cf. Banaji & Crowder 1989). Despite these difficulties, however, some attempts have been made to develop wholistic correspondence measures in certain circumscribed domains, primarily in the area of memory for visual and spatial information (see, e.g., Allen et al. 1978; Hart 1979; 1981; Pick & Lockman 1981; Siegel 1981; Siegel & Schadler 1977; Waterman & Gordon 1984). Here, measures of overall correspondence have been based on pattern-matching techniques to compute the *goodness of fit* between a particular target stimulus and its reconstruction from memory.

Three illustrative procedures, which convey some of the distinctive aspects of correspondence-based memory assessment, will now be considered. All pertain to the measurement of distortion in mental maps.

Waterman and Gordon (1984) had subjects draw the map of Israel from memory, and the correspondence of each map to the actual map was assessed with respect to eight clearly identifiable geographic points appearing in all of the map reproductions. After applying transformations designed to neutralize differences in rotation, translation, and scale, an overall "distortion index" was computed for each map in terms of the squared deviations (distances) between corresponding points on output and criterion maps. This index was normalized by expressing the measured distortion as the proportion of the maximum distortion possible given the assessment procedure.

In a similar vein, Siegel (1981) presented subjects with a simulated "campus walk" slide show and then obtained ordinal distance ratings between nine landmarks by the method of multidimensional rank order (Subkoviak 1975). Using nonmetric multidimensional scaling techniques, the ratings from each subject were transformed into a one-dimensional, best-solution map, whose correspondence with the one-dimensional representation of the actual route was computed by a goodness-of-fit procedure that compares the coordinates of points between  $n$ -dimensional arrays without regard to array shrinkage, expansion, or rotation. Using the resulting congruence index, Siegel found, for example, better memory for scenes with high versus low landmark potential, and for routes that were viewed twice rather than once.

A third example comes from Hart (1979; 1981), who used a somewhat different approach in assessing children's memory for the layout of their home town. Hart obtained memory "reports" from the children by having them build scale models. In evaluating the accuracy of the models, he identified and scored individual constellations of local features, as well as the more global configurations. The average local correspondence score was then multiplied by the global configuration score in order to obtain an "integrated map score." Hart noted, however, that the models varied greatly in the extent of the mapped area, that is, in the amount of information reproduced. He

therefore multiplied the integrated score by an "extent of area" score, arriving at an overall "composite map score" for each child.

These modest efforts illustrate some of the distinctive aspects of the wholistic assessment of memory correspondence: Most prominently, the memory report in each case is being assessed in terms of its *overall fit* with a complex target stimulus. In evaluating such a fit, the experimenter must always (at least implicitly) specify which features of possible correspondence are relevant and which are to be ignored. Thus, in the examples cited above, the overall orientation and scale of the reproduced maps were neutralized before correspondence was assessed, whereas in other cases orientation may be treated as a dimension of interest (e.g., Tversky 1981). Also in some cases the completeness of the report (e.g., extent of mapped area, number of landmarks) may be important, and the correspondence measure must be adapted to take this aspect into account (e.g., Hart 1981), or else the extent of the memory output may be controlled by the experimenter (e.g., Siegel 1981). In other cases, however, the experimenter may be concerned only in evaluating the accuracy of the information contained in the *output* without regard to its completeness (e.g., Waterman & Gordon 1984; and sect. 4.2.2.2). Finally, the correspondence measure may need to be expressed in normalized units (e.g., Waterman & Gordon 1984) in order to facilitate comparisons across different tasks. In sum, unlike traditional quantity measures, which have been designed as all-purpose tools, wholistic correspondence measures must be based on specific assessment models tailored to particular purposes.

The task of deriving an overall correspondence measure can become even more complicated when attempting to assess the overall correspondence between a real-life event and its verbal reconstruction. Consider, as an extreme example, the derivation of a single quantitative score that reflects the faithfulness of each of the four reports in Kurosawa's (1951) celebrated movie, *Rashomon*. The problem stems from the fact that real-life events can submit to a multitude of different descriptions, each of which may be "accurate" depending upon the specific evaluative criteria employed (see McCauley 1988; Neisser 1981; 1986; 1988c; for a discussion of similar problems regarding the accuracy of social perception, see Funder 1987; Kruglanski 1989). Thus, in *Rashomon*, the four versions of the crime agreed in many details, but the "critical" aspects, particularly those open to subjective interpretation, differed drastically from report to report. Conversely, there are cases when a memory report may be inaccurate in reproducing the details, or even the gist of what occurred on a specific occasion, but still convey a "true" impression of what was happening at the time (e.g., John Dean's memory in Neisser 1981; see also Spence 1982).

Clearly, then, the development of an overall assessment model capable of dealing adequately with the richness of real-life situations is a formidable task. In complex cases, a tempting option is to rely on subjective global accuracy ratings (e.g., Larsen 1988; but see Neisser & Harsch, 1992, for reservations). Indeed, an ingenious variation on the use of subjective judgments has been devised for the assessment of memory for faces, by determining the proportion of correct target recognitions achieved by independent judges on the basis of the memory report alone (Ellis et al. 1975; Wells & Turtle 1988). This procedure has also been

used to compare the effectiveness of verbal descriptions versus photofit reconstructions (Christie & Ellis 1980).

Of course, other techniques might also be envisaged. In fact, it is odd that the problem of measuring accuracy has received so little attention from students of memory, compared, for instance, to the systematic analysis it has received from students of social perception and social judgment. In the latter domains, numerous papers have specifically addressed the conceptual and methodological issues that arise in measuring the overall accuracy and inaccuracy of interpersonal (and intrapersonal) judgments (e.g., Cronbach 1955; Funder 1987; Kenny 1991; Kenny & Albright 1987; Kruglanski 1989; Sulsky & Balzer 1988). This work should be of interest to memory researchers as well (see also Ross, in press, for an interesting social-psychological treatment of the issue of memory accuracy). Sulsky and Balzer (1988), for example, compared five different conceptual and operational definitions of judgmental accuracy, all of which involved comparing a subject's judgments to a set of criterion judgments. An even wider range of conceptions of judgmental accuracy is examined by Kruglanski (1989; and see Ross, in press).

Overcoming the conceptual and methodological hurdles that impede the development of wholistic correspondence measures poses a crucial challenge for memory research. In principle, such measures could supply the needed tools for a bona fide *psychology of memory correspondence*, to parallel the quantity-oriented tradition. For example, they could enable researchers to trace the course of forgetting over time in the sense of a reduction in the overall faithfulness of memory, to examine the effects of a variety of factors on the rate of such forgetting, to study individual and group differences in memory accuracy, to explore the effectiveness of different questioning procedures in improving the faithfulness of memory reconstructions, and so forth. This is the type of approach that would seem to follow most naturally from many discussions of everyday memory, as well as from the correspondence metaphor itself (see further discussion in sect. 6). Because such data would undoubtedly be of great value, their absence from the memory literature is conspicuous. Greater efforts in this direction are certainly called for.

#### 4.2. The analytic evaluation of memory correspondence

The second, more common, approach to the evaluation of memory correspondence is the analytic approach. Indeed, in light of the complexities described above, it is not surprising that most research on memory accuracy has shied away from a direct confrontation with the wholistic assessment of memory faithfulness. In the study of autobiographical memory, for instance, most researchers still treat memory "as if it were just a set of remembered concrete experiences" (Neisser 1988c, p. 356). Such a treatment, which follows from the storehouse metaphor, is less suited to the evaluation of memory correspondence. Nevertheless, many of the unique concerns of correspondence-oriented memory assessment can still be expressed even when the analytic approach is adopted. Thus, particularly in view of the widespread use of analytic memory measures, it is important to examine how accuracy-oriented and quantity-oriented memory assessment differs in the context of this approach.

As a framework for our analysis, we assume an item-based memory testing situation (see Puff 1982) in which the target information has been segmented into discrete units (items or propositions). Quantity-based and accuracy-based memory measures may then be distinguished as follows: *Quantity* measures are *input-bound*, assessing the likelihood of correctly remembering (recalling, recognizing, etc.) an input item. *Accuracy* measures, in contrast, focus on the *dependability* of the reported information. Hence, they are *output-bound*, reflecting the likelihood that a reported item is "correct," that is, corresponds to the input (see Koriat & Goldsmith 1994). Whereas input-bound measures are traditionally used to estimate the amount of stored information that can be recovered, output-bound measures, being conditional on the output, evaluate the accuracy of the information that is reported. The latter measures are of particular interest in situations such as eyewitness testimony, in which the dependability of reported information is often no less important than its amount (see, e.g., Deffenbacher 1988; 1991; Fisher et al. 1989; Hilgard & Loftus 1979; Loftus 1979a; Wells & Lindsay 1985; Wells & Loftus 1984). As we shall now show, in some cases the two types of measures yield equivalent scores, though each implies a different "attitude" in its interpretation, whereas in other cases they differ operationally as well.

**4.2.1. The intent to measure accuracy.** Consider a simple situation in which memory is tested for only one item of information using a *forced-choice* procedure. For example, in the well-known study by Loftus, Miller, and Burns (1978), subjects were required to decide whether the traffic sign in the witnessed event was a stop or a yield sign. In that study and others like it (e.g., Boon & Davies 1988; Wagenaar & Boer 1987), memory accuracy is assessed simply by noting whether the provided answer is correct or incorrect. This might be compared to the hypothetical case in which a studied list of paired associates is followed by a single probe and two alternative responses, e.g., SIGN - STOP/YIELD, and the intention is to assess memory quantity. Operationally, the two measures, accuracy and quantity, are equivalent; the difference between them is solely a matter of the experimenter's *intent*. Whereas in the former case the test is designed to examine whether the person's memory is a faithful reproduction of the witnessed event, in the latter case the intent has traditionally been to determine whether the designated item is still in store and accessible.

The same is also true when memory for a list of items is tested through a forced-choice procedure. For example, assuming that the memory for 20 items of information is tested, and 12 items are answered correctly, then the likelihood of correctly remembering each item (quantity) and the likelihood that each answer is correct (accuracy) are both .60. In general, then, when a forced-choice, item-based procedure is employed, the exact same test can be used as a measure of either accuracy or quantity, and will yield the same memory performance score regardless of which property is intended.

How can the researcher's intent be distinguished in such cases? The intent to measure accuracy rather than quantity is sometimes explicitly stated by the investigator, as is typically the case in eyewitness research. In other cases, however, it can only be inferred from a variety of cues that disclose the implicit treatment of the subject's responses.

For example, an analysis of memory errors often discloses a focus on accuracy. Also, asking subjects to report how confident they are in the answer they chose (see Lichtenstein et al. 1982; Nelson & Narens 1990) may imply that the subjects' responses are being treated as propositional statements having a truth value. In general, however, the differences can be quite subtle.

#### **4.2.2. Distinguishing accuracy measures operationally.**

Notwithstanding these subtleties, there are two conditions for which item-based accuracy and quantity measures differ operationally as well. The first is when the stimulus information solicited from the subject may be evaluated on a continuous dimension, and the second is when the option to reply is under the subject's control.

**4.2.2.1. Dimensional accuracy:** As noted earlier, storehouse-inspired quantity measures are typically based on some type of counting operation, with individual items scored in a dichotomous (present/absent or correct/incorrect) fashion. Accuracy, in contrast, is more graded in nature, admitting different degrees of deviation from veridicality for any continuous or ordered dimension. For instance, given that the height of a burglar was actually 5 feet 8 inches, a report of 6 feet is clearly less accurate than a report of 5 feet 10 inches.

It is easy to overlook the fact that dimensional accuracy assessment is quite foreign to the storehouse metaphor, and implies a correspondence metaphor instead. Indeed, the measurement techniques themselves are often borrowed from the study of perception and are most readily applied when the memory target is the value of some perceptual attribute. Thus, many studies of memory for visual form and spatial information have evaluated subjects' memory reports with respect to such biases as increased closure and symmetry, changes in orientation, angular and radial deviation, and so on (e.g., Bartlett 1932; Byrne 1979; Goldmeier 1982; Huttenlocher et al. 1991; McNamara 1986; Nelson & Chaiklin 1980; Riley 1962; Tversky 1981; Tversky & Schiano 1989). Also, studies on the "psychophysics of memory" (e.g., Algom & Cain 1991; Algom et al. 1985; Kerst & Howard 1978; Moyer 1973; Moyer & Dumais 1978), have shown, for example, how "memory scale values map onto their physical referents via the same functional relation (power transform) as perceptual scale values do" (Algom et al. 1985, p. 468). More generally, dimensional accuracy has been investigated with regard to memory for the date and time of past events (e.g., Baddeley et al. 1978; Huttenlocher et al. 1990; 1988; Linton 1975; Loftus & Marburger 1983; White 1982) and for a variety of other variables ranging from SAT scores (Bahrick et al. 1993) to the height of Mt. Everest (Yaniv & Foster 1995).

Dimensional accuracy assessment was already implied in the preceding discussion of the wholistic assessment of memory correspondence. However, when only a single attribute is of concern, many of the problems involved in integrating information across dimensions can be avoided while still adhering to a correspondence type of measurement.

#### **4.2.2.2. The option of free report: input-bound versus output-bound measures:**

The second case in which accuracy and quantity measures differ operationally applies to the more standard item-based procedures, in which memory is tested for a set of dichotomously scored items. As indicated above, with such procedures, a forced-

choice memory report yields equivalent performance whether it is evaluated for quantity or accuracy. This evaluation can differ substantially, however, under *free-report* conditions, where subjects are free to volunteer or withhold information.

Consider, for example, an eyewitness who is asked to remember which people she saw at the scene of a crime and reports that she saw A, B, and C. If A, B, and C were indeed present, then this testimony is entirely *accurate*. The fact that other people, D and E, were also present but were not reported by the witness will not detract from the (output-bound) accuracy of the information that was provided. In contrast, construed as a free-recall task intended to tap the (input-bound) amount of information that can be reproduced, reporting only three people out of five will obviously count against the reporter. Thus, for quantity-based measures, omission errors (failing to report an item of information, i.e., loss of information) are the more serious errors, whereas for accuracy measures, commission errors (falsely reporting something that did not occur, i.e., loss of dependability) are critical and omissions may be ignored.

In general, input-bound and output-bound measures will be equivalent when the output list (e.g., people reported as being present at the scene of a crime) is the same length as the input list (people actually present), and such is the case when a forced-report procedure is used. Under free-report conditions, however, the option to reply is *controlled by the subject*, and therefore input-bound and output-bound measures may be expected to differ. Thus, the operational distinction between output-bound accuracy (i.e., "dependability") and input-bound quantity is applicable in the context of the standard item-based approach, but only when the subject is free to decide whether to volunteer or withhold specific items of information.

The role of *report option* in differentiating accuracy-based and quantity-based memory measures illustrates how a concern with memory correspondence may bring to the fore issues that are less intrinsic to a storehouse framework - in this case, the active role of the rememberer in controlling his or her memory output. Indeed, the issue of subject control may be helpful in elucidating the proposed relationship between memory metaphors and the methods, content, and context of memory research. For example, as will be discussed later (see sect. 6.1), not only are conceptual metaphors instrumental in dictating the methods by which memory should be studied and evaluated, but there is also a reciprocal relationship, in which the methods may illuminate and elaborate certain aspects of the metaphor. Thus, although subject control over memory performance is perhaps an optional aspect of the correspondence metaphor (cf. the "picture" or "copy" metaphors argued against by Neisser, 1967), as a practical matter, such control can have a substantial impact on measures of memory accuracy (particularly in naturalistic settings; see the following section). This, in turn, calls for a more "active" correspondence conception, leading to a focus on further, substantive issues that might otherwise be overlooked.

In the following section, we summarize experimental work that we have done that demonstrates more concretely the importance of the correspondence-storehouse distinction in item-based memory research. Although the item-based approach is restrictive for the evaluation of memory correspondence, it has the advantage of allowing some of the important features that distinguish the accuracy-

oriented and quantity-oriented approaches to memory (e.g., differences in subject control) to be directly compared within a common framework.

## 5. The accuracy-oriented and quantity-oriented approaches in item-based research

In the work reviewed below, we first show how interest in the two memory properties, accuracy and quantity, tends to be confounded not only with the contrast between the everyday and laboratory approaches, but also with other important dimensions of memory assessment. This analysis will help clarify some apparent inconsistencies that arise in comparing accuracy-oriented, naturalistic findings and quantity-oriented, laboratory findings. We then focus on the issue of subject control over memory reporting, its unique role in the accuracy-oriented study of memory, and the challenges that it provides for memory research and assessment. Finally, we consider some of the implications of this work for the real-life/laboratory controversy.

### 5.1. Dimensions of memory assessment

Our experimental work (Koriat & Goldsmith 1994; in press b) was originally motivated by some apparent discrepancies between findings obtained in the laboratory and those obtained in naturalistic contexts. In attempting to resolve these discrepancies, we identified four dimensions of memory assessment that tend to be confounded in the reality of memory research. The first dimension is the *context of inquiry* dimension, that is, laboratory versus real life. The second dimension is the *memory property* of interest, quantity versus accuracy. As discussed throughout this paper, much of the research carried out under the banner of everyday memory reveals a special concern with the accuracy or dependability of memory, in contrast to the predominant focus on memory quantity in traditional, laboratory-based research.

The third dimension is *report option*, or the extent to which subjects are allowed to control their memory reporting by choosing which items of information to volunteer and which to withhold. As noted above, this dimension is crucial for the operational distinction between accuracy and quantity memory measures. Moreover, subject control over memory reporting is also generally confounded with the context-of-inquiry dimension: In traditional laboratory research, perhaps because of the high premium placed on experimental control (Banaji & Crowder 1989; Nelson & Narens 1994), subjects are generally given relatively little control over their memory reporting. In contrast, in naturalistic research situations, as in everyday life, people are typically allowed much more freedom in choosing what aspects of the event to relate, which to play down or ignore entirely, what perspective to adopt, how hard to try and get the details right, and so forth (see, e.g., Fisher & Geiselman 1992; Hilgard & Loftus 1979; Neisser 1981; 1988b; Nigro & Neisser 1983). Such control can substantially enhance output-bound accuracy, and indeed, in everyday contexts, for instance in the type of free-narrative reporting commonly used to obtain accurate reports from witnesses (see Hilgard & Loftus 1979), it may be employed precisely toward that end.

The fourth dimension is *test format*. Test format refers to

whether subjects *produce* their own answers or instead must *select* or *recognize* a response from among those provided by the experimenter.<sup>8</sup> As will be discussed later, this dimension, too, is often confounded with the previously mentioned factors, complicating further the comparison of results between laboratory and everyday settings.

A study reported by Neisser (1988b) illustrates some of the issues that stem from these confoundings. Neisser examined memory for real-life events that took place during the course of a seminar he taught. Memory was assessed using either a cued recall or a multiple-choice recognition procedure. Neisser found recall memory to be much more accurate than recognition memory, and pointed out that such a finding might come as a surprise to traditional memory researchers, who are accustomed to the general superiority of recognition memory found in laboratory studies.

Neisser's finding brings to the fore some of the potential sources of confusion in the comparison of findings obtained in naturalistic and laboratory research contexts, and illustrates what we have called the *recall-recognition paradox* (Koriat & Goldsmith 1994). On the one hand, the established wisdom in eyewitness research holds that testing procedures involving recognition or directed questioning can have "contaminating" effects on memory (see, e.g., Brown et al. 1977; Gorenstein & Ellsworth 1980; Hilgard & Loftus 1979; Loftus 1979a, 1979b; Loftus & Hoffman 1989). In fact, the general recommendation is to elicit information initially in a free-narrative format before moving on to directed questioning, and even then, to place greater faith in the former (see Fisher et al. 1987; Hilgard & Loftus 1979). On the other hand, however, this body of evidence stands in seeming defiance of the well-established superiority of recognition over recall memory in traditional, list-learning laboratory experiments (e.g., Brown 1976; Shepard 1967; but see Tulving & Thomson 1973). Thus this discrepancy - and Neisser's finding - could be taken as yet one more example of the importance of factors specific to the context of inquiry, that is, real-life versus laboratory contexts, supporting the claim that memory behaves differently in the two settings (e.g., Baker-Ward et al. 1993; Conway 1991; 1993).

The discrepant findings, however, may also be related to the two different memory properties, accuracy and quantity: In Neisser's (1988b) study, for instance, as in many naturalistic studies, the focus is on memory accuracy, whereas traditional memory research has focused almost invariably on memory quantity. Thus Neisser's recall subjects were more accurate than the recognition subjects in the sense that what they reported was almost never wrong, but as Neisser also pointed out, they did not provide much information either. Such findings could therefore reflect an interaction between memory property and test format that would be obtained, perhaps, regardless of the research context: Recognition yields better *quantity* performance than recall testing, but recall yields better *accuracy* (see also Hilgard & Loftus 1979; Lipton 1977).

An additional complication, however, stems from the common confounding of test format and report option. This confounding is evident in the reality of both naturalistic and laboratory research: In naturalistic research, for instance, free-narrative reporting not only guards against "leading" or contaminating information (a test-format variable), it also allows the witness the freedom to choose what information

to report, and at what level of generality. Directed questioning, on the other hand, often involves explicit or implicit demands that an answer be provided. Similarly, traditional item-based laboratory research almost invariably implements recognition testing as *forced* recognition in two distinct respects: Not only are subjects confined to the alternatives presented (test format), they are also forced to answer each and every item (report option). In contrast, recall testing typically allows subjects the freedom to decide both how and whether to report what they remember.

A further possibility, then, is that the recall-recognition paradox actually reflects an interaction between report option and memory property. Indeed, Neisser (1988b) pointed out that his recall subjects seemed to achieve greater accuracy by providing fewer answers. In addition, they might also have utilized a different aspect of report option to boost their accuracy: control over the "grain size" or generality of their responses (cf. Yaniv & Foster 1990; 1995). Clearly, the correspondence between memory reports and past events can improve when the answers are more general and less detailed. Thus, Neisser observed that his recall subjects tended to choose "a level of generality at which they were not mistaken" (1988b, p. 553).

## 5.2. Disentangling memory property, report option, and test format within a laboratory research context

The confoundings discussed in the previous section led us to propose a three-factor classification of item-based memory assessment methods in terms of memory property, report option, and test format (see Fig. 1). In proposing this scheme (Koriat & Goldsmith 1994a), we tried to show how it could serve as a guide in disentangling some of the empirical confusions discussed above, and also provide an integrative framework that might be exploited in future research. In what follows, we sketch some findings obtained within this framework (Koriat & Goldsmith 1994; in press b), focusing on their implications for the distinction between the accuracy-oriented and quantity-oriented approaches to memory.

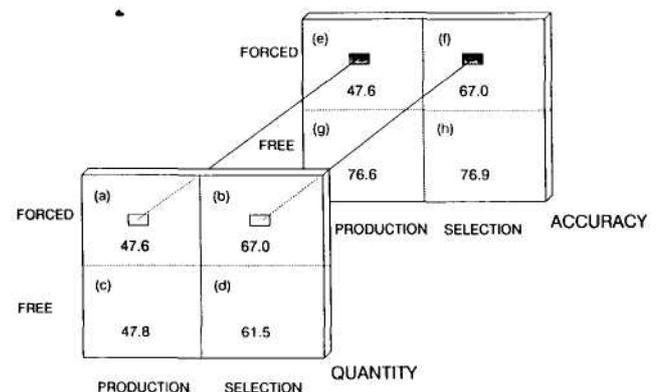


Figure 1. A three-way classification of item-based assessment methods in terms of memory property (quantity vs. accuracy), report option (free vs. forced), and test format (production vs. selection). Presented in each cell is the corresponding performance measure (mean input-bound quantity percentage or mean output-bound accuracy percentage) obtained for each method in Koriat and Goldsmith (1994, Experiment 1). The connecting lines indicate that the forced-report quantity and accuracy measures are operationally equivalent.

In one experiment (Koriat & Goldsmith 1994, Experiment 1), subjects answered 60 general-knowledge questions in either a recall or a multiple-choice recognition format (all items required a one-word answer in order to equate the "grain size" of the answers across the two test formats). In addition to the standard tests of free recall and forced-choice recognition, however, two relatively uncommon procedures were added: forced recall (requiring subjects to respond to all questions) and free recognition (permitting subjects to skip items). In this design, then, test format and report option were orthogonally manipulated. A payoff schedule provided all subjects with a common performance incentive, essentially rewarding them for each correct answer, but penalizing them by an equal amount for each incorrect answer. Performance was scored both for quantity (input-bound percent correct) and for accuracy (output-bound percent correct).

The results, superimposed on the classification in Figure 1, disclose several trends: When comparing the standard memory measures, free recall and forced recognition, our results replicated the pattern implicated in the recall-recognition paradox: Recall was indeed superior to recognition on the accuracy measure (g vs. f), but recognition was superior to recall on the quantity measure (b vs. c). However, examination of the remaining means indicates that although memory *quantity* performance does vary with *test format*, recognition yielding better quantity scores than recall (b vs. a and d vs. c), it is *report option* that is critical for memory *accuracy*: The option of free report increased accuracy performance for both recall and recognition testing (g vs. e and h vs. f). In fact, under free-report conditions (in which memory accuracy and quantity measures can be operationally distinguished), test format had no effect at all on memory accuracy: Given equal opportunity to screen their answers, the recall and recognition subjects achieved virtually identical accuracy scores (g vs. h)!

This basic pattern was replicated in several further experiments, employing both list-learning (episodic) and general knowledge (semantic) memory tasks (Koriat & Goldsmith 1994; in press b). What are the implications of these findings? First and foremost, the results demonstrate the importance of distinguishing between the two memory properties, memory quantity and memory accuracy. These properties were found to be dissociable: Test format affected quantity performance but not accuracy, whereas report option affected accuracy but not quantity. Thus, it is clear that one can neither compare nor investigate "memory performance" without specifying the memory property of interest and the particular testing conditions under which it is evaluated. Second, the results question the general belief among everyday memory researchers that recognition testing is necessarily detrimental to memory accuracy. Although this may often be the case, it could be due primarily - perhaps entirely - to the typical confounding of recognition testing with forced report. Third, because the superior accuracy of free recall over forced recognition characteristic of naturalistic research was obtained in these experiments within a typical laboratory setting, it would appear that at least some of the underlying dynamics are not uniquely tied to real-life contexts (see further discussion in sect. 5.4).

More generally, the results highlight the utility of the distinction between the accuracy-oriented and quantity-oriented approaches to memory assessment even within the

standard item-based research framework. Beyond helping to unravel some of the confusions involving memory properties and the other factors, the results underscore the need for a more careful consideration of some of the unique emphases of the accuracy-oriented approach, most notably the active role of the subject in controlling his or her output-bound memory correspondence. We shall now consider the ramifications of such control for the correspondence-oriented study of memory.

### 5.3. The strategic regulation of memory performance

It is perhaps not coincidental that the issue of subject control figures prominently, both when comparing everyday and laboratory research in general and when contrasting the accuracy-oriented and quantity-oriented approaches to memory in specific. As noted earlier, the investigation of real-life memory phenomena sometimes requires a compromise between the desire for strict experimental control, on the one hand, and the wish to remain true to the natural dynamics of the memory phenomena being investigated, on the other (see Gruneberg & Morris 1992). Hence, there seems to be a greater willingness among students of real-life memory to allow subjects control over their memory reporting, as is seen, for instance, in the use of free-narrative and other open-ended questioning techniques which are seldom employed in traditional memory research (see Hilgard & Loftus 1979).

In parallel, at the empirical level, the effects of subject control on memory performance also appear to differ markedly between the quantity-oriented and accuracy-oriented research contexts. On the one hand, quantity-oriented memory research suggests that subjects have little control over their memory performance: First, subjects cannot improve their memory quantity scores when given incentives to do so (e.g., Nilsson 1987; Weiner 1966a; 1966b). Second, encouraging or even forcing recall subjects to produce more items (by relaxing their response criteria) does not seem to improve their memory quantity scores much or at all beyond what is obtained under standard instructions (e.g., Bousfield & Rosner 1970; Erdelyi et al 1989; Roediger & Payne 1985; Roediger, Srinivas et al. 1989). On the other hand, in sharp contrast to these results, accuracy-based findings (e.g., Barnes et al. 1995; Koriat & Goldsmith 1994; in press b) indicate that *accuracy performance is under strategic control*: Not only can subjects improve their accuracy considerably when offered the option of free report (as discussed above), they can also increase their accuracy even further when given stronger incentives to do so. For example, in one experiment (Koriat & Goldsmith 1994, Experiment 3), we used the same free-report procedure described earlier, but this time subjects sacrificed all winnings if they volunteered even a single incorrect answer. Accuracy increased substantially compared to our earlier experiment, averaging over 90% for both recall and recognition. (Fully one quarter of the subjects were successful in achieving 100% accuracy!) This improvement, however, was attained at a cost in quantity performance (about a 25% reduction for both recall and recognition). Similar results were obtained using a 10:1 penalty-to-bonus payoff ratio (Koriat & Goldsmith, in press b).

Clearly, then, subject control over memory reporting should be of special concern in correspondence-oriented memory research. In recounting experiences, people can

apparently regulate their reporting so that accuracy is enhanced (among other goals; see Neisser 1988b; Ross & Buehler 1994): They may report only information about which they are confident (Barnes et al. 1995; Koriat & Goldsmith 1994; in press b), or adopt a level of generality at which they are not likely to be wrong (Neisser 1988b; Yaniv & Foster 1990; 1995). This creates two fundamental challenges for student's of memory correspondence: First, how can subject-controlled regulatory processes be made amenable to experimenter-controlled, scientific investigation? Second, given that memory correspondence is under the control of the subject, how can such control be accommodated by our methods of memory assessment?

**5.3.1. Investigating subject control over memory reporting.** As indicated above, in our work we attempted to address these questions by focusing on one specific type of subject regulation - the withholding or volunteering of particular items of information in free-report situations. As a framework for investigating such regulation, we put forward a model of monitoring and control processes that merges ideas from signal-detection theory with ideas from metamemory research (Koriat & Goldsmith, in press b). We assume that, when attempting to recount a past event, people monitor the subjective likelihood that an item of information that comes to mind is correct and then apply a control threshold to the monitoring output in order to decide whether to volunteer that item or not. The setting of the control threshold is assumed to depend on the relative utility of providing complete versus accurate information. Several results supported the model: First, the tendency to report an answer increased strongly with increasing confidence in the correctness of the answer. Second, subjects given a high accuracy incentive (a 10:1 penalty-to-bonus ratio) were more selective in their reporting, adopting a stricter criterion than subjects given a more moderate incentive (a 1:1 ratio). Third, by using these monitoring and control processes, subjects were able to enhance their memory accuracy under free-report conditions, but the accuracy improvement generally came at the expense of quantity performance.

These results can help shed light on some of the mechanisms underlying the strategic control of memory accuracy. In particular, one aspect of the work is especially useful in demonstrating how the correspondence-oriented concern with memory accuracy leads to very different research emphases than the traditional concern with memory quantity: Despite the intuitive appeal of a criterion-based quantity-accuracy tradeoff (implied by signal-detection theory), our work shows that neither the accuracy advantage that typically derives from subject control over memory reporting nor the quantity costs of such control are inevitable. Rather, these effects depend critically on both *accuracy motivation* and *monitoring effectiveness*. These two factors have been virtually ignored in traditional, quantity-oriented memory assessment (which might in fact account for Roediger, Srinivas, and Waddil's [1989] observation that a recall-criterion effect on quantity performance is "intuitive, but remarkably little evidence for it exists" [p. 255]). Let us, then, expand briefly on these two factors.

First consider accuracy motivation. In previous, quantity-oriented investigations of motivational effects on memory (e.g., Nilsson 1987; Weiner 1966a; 1966b), the incentive manipulations were explicitly designed to increase memory

quantity performance, and null effects were taken to imply that motivation "does not affect memory performance" (Nilsson 1987, p. 187). By contrast, our results indicate that, when the focus is on *accuracy* performance, the effects of *accuracy motivation* on both accuracy and quantity measures can be substantial. Similarly, in the demonstrations of null or very small effects of recall criterion on memory quantity performance mentioned earlier, there was no special motivation for accuracy in either the experimental or control conditions (forced-report vs. standard free-recall instructions). Had those studies, like our experiments, included a condition with a strong incentive for accuracy, they too would undoubtedly have found the ensuing changes in criterion level to yield substantial effects on both quantity performance and accuracy performance.<sup>9</sup> However, in those studies, accuracy performance, and hence accuracy motivation, were of no direct interest.

The second factor that should be of special concern in accuracy-oriented research is the effectiveness of subjects' memory monitoring. Although this factor has attracted much attention among students of metacognition (see Metcalfe & Shimamura 1994), its performance consequences have received relatively little attention (but see Barnes et al. 1995; Bjork 1994; Metcalfe 1993; Nelson & Narens 1994). In the context of our research, monitoring effectiveness refers to the correspondence between the subjects' confidence regarding a candidate answer and the actual probability that the answer is correct.<sup>10</sup> It is important to note that this factor is distinct from the amount of information retained.<sup>11</sup> To illustrate, consider a relatively difficult memory test, for which a subject fails to remember the answers to many items. Even though memory retention (and hence quantity performance) may be poor, the subjects' *monitoring* of the correctness of their answers could still be perfect. In that case, the option of free report would allow him or her to volunteer only (the few) correct answers, achieving perfect accuracy with no tradeoff. On the other hand, the subjects' monitoring might be very poor as well, in which case utilizing the option of free report should not enhance his or her accuracy much or at all, and would only reduce quantity performance.<sup>12</sup>

The important implication of this analysis is that monitoring effectiveness can influence memory performance, particularly memory accuracy, independent of what might be called memory "retention." Thus, in one experiment (Koriat & Goldsmith, in press b, Experiment 2) we manipulated monitoring effectiveness by using two different sets of general-knowledge items: One set (the "poor" monitoring condition) consisted of items for which the subjects' confidence judgments were expected to be generally uncorrelated with the correctness of their answers (see Fischhoff et al. 1977; Gigerenzer et al. 1991; Koriat 1995), whereas the other set (the "good" monitoring condition) consisted of more typical items, for which the subjects' monitoring was expected to be more effective. As predicted, although the two sets were matched on retention, as indexed by *forced-report* quantity performance, the good monitoring condition allowed subjects to attain a far superior joint level of *free-report* accuracy and quantity performance: much better accuracy performance was achieved while maintaining equivalent quantity performance, compared to the poor monitoring condition (see Fig. 2).

Our work, then, indicates that free-report memory measures tap the operation of memory components that are not

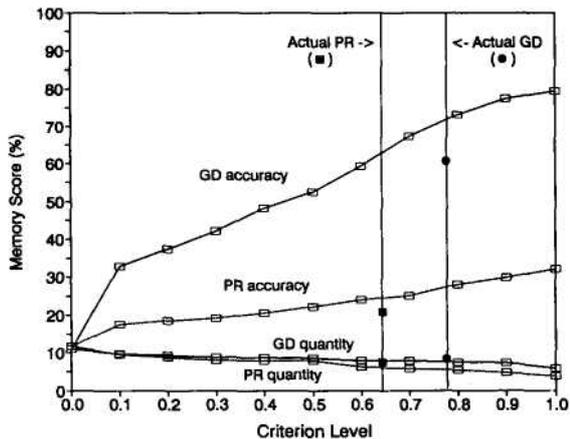


Figure 2. The memory performance effects of good versus poor monitoring. Expected memory accuracy and memory quantity scores (output-bound and input-bound percent correct, respectively) are plotted as a function of criterion level for the poor (PR) and matched-good (GD) monitoring conditions in Koriat and Goldsmith (in press b, Experiment 2). The plots are derived by incrementally applying a control threshold to each subjects forced-report responses on the basis of the confidence associated with each answer (forced-report performance is plotted at criterion = 0). In addition, each subject's actual free-report control decisions (to volunteer or withhold each candidate answer) are used to estimate the criterion level adopted by that subject and to calculate actual free-report performance. The mean free-report accuracy and quantity scores for each monitoring condition are plotted as bullets above the mean criterion estimate for that condition. (For more details, see Koriat & Goldsmith, in press b)

disclosed by forced-report measures, and that these components have a critical role in the strategic regulation of memory accuracy. In particular, accuracy motivation and monitoring effectiveness emerge as crucial factors in determining memory accuracy under free-report conditions and in modulating the rate of the quantity-accuracy tradeoff. Thus, this work motivates a greater concern with both the determinants and the accuracy of metacognitive processes as they affect memory performance (see also Nelson & Narens 1994).

**5.3.2. Incorporating subject control into the assessment of memory correspondence.** The issue of subject control also presents a dilemma with regard to memory assessment: How can we sensibly assess a person's memory for an event if memory performance, particularly memory accuracy, is under the person's control? This issue is not just methodological, but also metatheoretical: the question is whether intervening activities on the part of the subject, such as deciding to volunteer or withhold information, are to be conceived as operations that are superimposed on memory (see Klatzky & Erdelyi 1985; Lockhart & Craik 1990) or rather as being part and parcel of memory itself (see Nelson & Narens 1994; Tulving 1983).

As alluded to earlier, traditional storehouse-guided research has tended to treat subject control as a nuisance variable that should be eliminated or partialled out in order to achieve a pure measure of "true" memory. Indeed, Nelson and Narens (1994) note that "ironically, although the self-directed [subject-controlled] processes are not explicitly acknowledged in most theories of memory, there is an implicit acknowledgment on the part of investigators

concerning the importance of such processes. The evidence for this is that investigators go to such great lengths to design experiments that eliminate or hold those self-directed processes constant via experimental control!" (p. 8).

Even when self-directed processes are not experimentally controlled, an attempt is often made to partial out their effects in order to derive a pure measure of memory. For instance, the signal-detection measure of true memory ( $d'$ ) is designed to provide an estimate of memory strength (or sensitivity) that is "unbiased" by variation in  $\beta$  (commonly referred to as "response bias").<sup>13</sup> In this sense, signal-detection methods are often used like other techniques that correct for the effects of guessing (e.g., Budescu & Bar-Hillel 1993; Cronbach 1984; see also Koriat & Goldsmith, 1994, and further discussion below).

Such a treatment, however, would seem unsuitable for the evaluation of memory under the correspondence metaphor. Because subject-controlled metamemory processes actually constitute an important means of managing one's memory correspondence, they cannot simply be avoided or partialled out. On the contrary, when the researcher is explicitly concerned with the faithfulness of memory, and in particular with the dependability of memory reports in real-life settings, it would seem imperative to treat the ongoing regulation of memory performance as an intrinsic aspect of memory functioning (Neisser 1988b; see also Barnes et al. 1994; Koriat & Goldsmith 1994; in press b; Metcalfe & Shimamura 1994; Nelson & Narens 1990; 1994). An important challenge, then, is to develop ways of making the contribution of metamemory processes explicit in the evaluation of memory performance.

One method that we have proposed was, in fact, already illustrated at the group level in Figure 2. Rather than seek a single-point estimate of "true" memory, this method incorporates metamemory processes into memory assessment by charting memory performance profiles that take retention, monitoring, and control into account (Koriat & Goldsmith, in press b). This approach resembles that of plotting memory operating characteristic (MOC) curves using signal-detection techniques (but see below). Like an MOC curve, the proposed *quantity-accuracy profile* (QAP) describes the joint levels of quantity and accuracy performance that can potentially be achieved under different conditions.

Like MOC curves, QAPs also can be plotted at the individual level. Consider, for example, the two QAPs depicted in Figure 3, which were computed for two selected recall subjects from our recent study (Koriat & Goldsmith, in press b). If we were to look only at forced-report performance (criterion = 0) as a point-estimate of memory retention, subject B's performance would clearly be better than A's. Similarly, if we were to look only at the subjects' actual free-report memory scores (ignoring or perhaps overlooking important differences in accuracy motivation), B would be seen to achieve about equal memory quantity, but far superior accuracy than A. The profiles, however, offer much more than this. First of all, the QAP makes quite clear that the two subjects adopted different criteria in controlling their actual free-report responding. (In fact, Subject A responded under a moderate accuracy incentive, whereas B responded under a strong accuracy incentive.) Second, looking at potential memory performance, not only is B's potential quantity performance superior to A's across the range of criterion levels, but B's better monitoring effectiveness (.87 vs. .64 on the ANDI measure; see Yaniv et

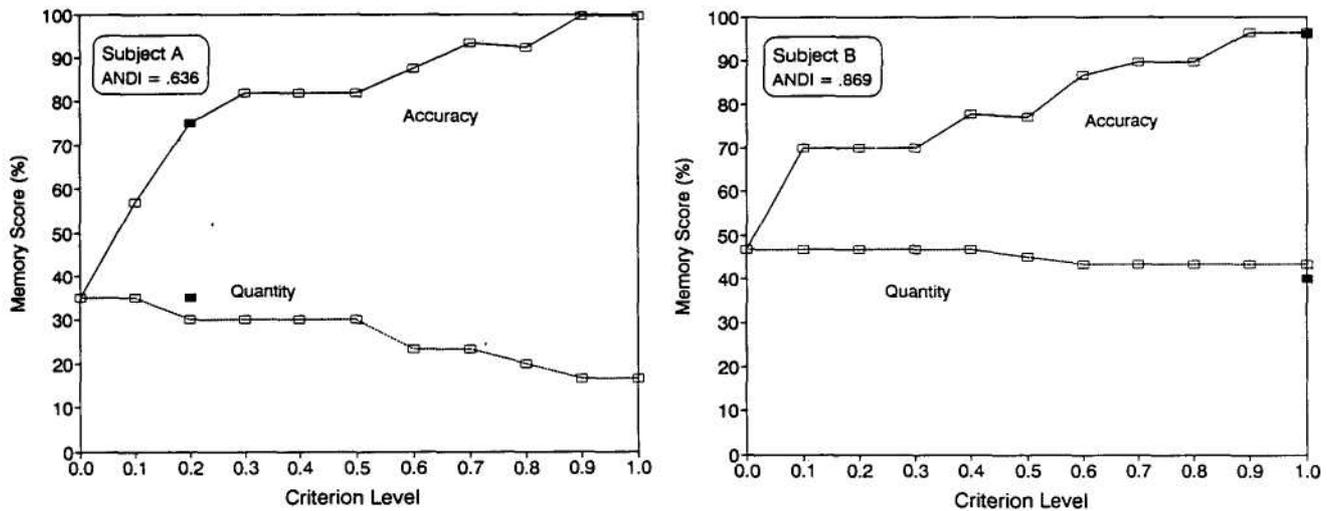


Figure 3. Quantity-accuracy profiles (QAPs) for two selected recall subjects from Koriat and Goldsmith (in press b, Experiment 1). For each subject, the profile plots expected accuracy and quantity performance (output-bound and input-bound percent correct) as a function of criterion level, derived on the basis of the subject's forced-report performance and the confidence associated with each answer. In addition, the actual accuracy and quantity scores achieved under free-report testing are plotted as bullets above the estimated criterion setting used by the subject. The ANDI measure of monitoring resolution (Yaniv et al. 1991) is also presented.

al. 1991) allows a high level of accuracy to be achieved at virtually no cost to the number of correct answers provided. On the other hand, A is potentially able to achieve 100% accuracy (though at a substantial quantity cost) when strongly motivated for accuracy, while B's maximal accuracy falls somewhat short of this. Ultimately, then, in order to evaluate the relative effectiveness of A's and B's memories, we will need to take into account functional considerations pertaining to the circumstances of the memory report. For instance, as a key witness in a capital trial, we might actually prefer A's memory, because of the very high premium placed on memory accuracy in such situations.

Compared to the standard point measures of memory performance, the derivation of quantity-accuracy profiles allows a more global evaluation of potential memory performance in terms of both accuracy and completeness. This approach also illustrates one way in which memory assessment can accommodate the contribution of subject-controlled metamemory processes to overt memory performance. Indeed, QAPs may be used to separate the sources of individual or group differences (e.g., developmental changes) and the effects of different manipulations on memory retention, monitoring, and control in a manner similar to the way in which signal-detection methods allow one to distinguish differential effects on  $d'$  and  $c$  (see Koriat & Goldsmith, in press b).

This brings us to an important point, however, which we now address. Clearly, there is an overall resemblance between our proposed framework and the signal-detection approach to memory measurement (see, e.g., Banks 1970; Bernbach 1967; Kintsch 1967; Klatzky & Erdelyi 1985; Lockhart & Murdock 1970; Murdock 1966; Norman & Wickelgren 1969). Signal detection is often thought to represent an "accuracy-oriented" approach to memory, and it indeed raises many of the same issues brought out here. Yet, despite the apparent similarities, there are several fundamental differences that should be emphasized.

First and foremost, the signal-detection methodology cannot be applied at all to free-report situations, precisely because in such situations subjects have the option to

decide whether to volunteer or to withhold information (see Lockhart & Murdock 1970). In fact, the response criterion ( $P$ ) addressed by the signal-detection methodology is not whether to respond or not, but rather whether to respond "old" (studied) or "new" (foil) to each and every item (under forced report). Thus, while the signal-detection approach has contributed greatly to a consideration of the role of decision processes in forced-recognition memory, it actually has little to say regarding the accuracy of a person's freely reported remembrances.<sup>14</sup>

Second, the signal-detection methods do not distinguish between "retention" and "monitoring" (see Koriat & Goldsmith, in press b; Lockhart & Murdock 1970). For instance, in the forced-report old/new paradigm to which signal-detection methods are typically applied, "control" is isolated in terms of the parameter ( $c$ ), yet "retention" (overall memory strength) and "monitoring effectiveness" (the extent to which the subject's confidence distinguishes "old" from "new" items) are operationally equivalent: Both are equally valid interpretations of  $d'$  (see, e.g., Banks 1970; Lockhart & Murdock 1970).<sup>15</sup> By contrast, as discussed above, in our approach to free-report performance, these latter two aspects (as well as control) may be evaluated independently: one may have good monitoring resolution, yet very poor retention, or vice versa.

Third, although the logic of signal detection can be extended to free-report tasks (see Klatzky & Erdelyi 1985; Koriat & Goldsmith, in press b), the motivation for doing so has generally been to control for criterion effects (differences in accuracy) when comparing quantity measures (e.g., by using forced-recall procedures; Erdelyi & Becker 1974), rather than to measure memory accuracy as a property of interest in its own right (see also Koriat & Goldsmith 1994). Thus, for instance, in discussing the possible implications of a null effect of recall criterion on memory quantity performance, Erdelyi et al. (1989) observe:

If response bias fails to affect recall performance level, forced-recall procedures for controlling shifts in productivity (e.g., Erdelyi & Becker, 1974), which subjects find tedious, can be dispensed with in favor of standard free recall. Of more impor-

tance, the difficulty of applying forced recall to complex, ecologically more valid materials than stimulus lists - such as prose passages or actual real-life events - becomes moot, and recall performance, irrespective of intrusion levels, could be trusted to reflect *true recall level*. Would that this were so, for no proven methodologies have been worked out for controlling response bias in the recall of complex stimuli for which forced recall is unwieldy or inapplicable, (p. 246, emphasis added) As

already noted, this same quantity-oriented attitude is also generally evidenced in the evaluation of forced-report recognition memory, in which case the signal-detection methodology can actually be applied.

Overall, then, our work demonstrates how some rudimentary aspects of subject control in situations of free-memory reporting can be experimentally studied and taken into account in the assessment of memory correspondence. It also illustrates how, even within an item-based framework, the accuracy-oriented approach to memory brings to the fore questions that might be neglected when the focus is strictly on memory quantity. Of course, in everyday life, people have many more means available to manage their memory correspondence than just the simple option of volunteering or withholding specific pieces of information (cf. John Deans memory in Neisser 1981). Thus, a better understanding of the faithfulness of memory in real-life contexts will require greater efforts to bring these other aspects of subject control under systematic investigation.

#### **5.4. Implications for naturalistic and laboratory-based research**

In concluding this section, let us reconsider how the research discussed above bears upon the context-of-inquiry issue (i.e., naturalistic versus laboratory contexts).

As noted earlier, many everyday memory researchers stress factors related to the context of research in explaining differences in memory performance (e.g., Baker-Ward et al. 1993; Fivush 1993; Conway 1991; 1993; Neisser 1988b). These factors generally have to do with the functional role of remembering in naturalistic settings. Neisser (1988b), for instance, argued that in real-life situations, forces operating at the time of remembering may have much more impact than has been acknowledged in traditional laboratory-based research. In his words, "we must take into account not only the stimuli present at retrieval but the reason for retrieval; the theory we require will have to deal with persons, motives, and social situations" (1988b, p. 553). These remarks suggest the possibility of inherent differences in the dynamics of remembering between everyday and laboratory contexts, which may limit the generalizability of results across the two contexts (see also Conway 1993). Thus, for example, in noting the relatively high accuracy of eyewitness memory in naturalistic field studies, Fisher et al. (1989) observe:

It is interesting that the accuracy-corroboration rates in the three field studies of eyewitness memory were considerably higher than their laboratory counterparts. If this difference between laboratory and field studies continues to appear, one may question the validity of describing in court the accuracy rates found in the laboratory as evidence of the general unreliability of eyewitness testimony in field cases, (p. 4, note 6) The important question, however, is whether some of the sources of the differences in memory performance between naturalistic and laboratory contexts can be identified.

Although context of inquiry was not manipulated in our

research, the results nevertheless suggest some factors that might underlie discrepancies between naturalistic and laboratory findings. The most crucial factor, of course, is the memory property being assessed — quantity versus accuracy. Because most laboratory research has concerned itself with memory quantity, the focus of everyday memory research on accuracy could give the impression that memory performance is different in naturalistic settings than in the laboratory. Indeed, as discussed earlier, this factor partly accounts for the apparent discrepancy between the established superiority of recall over recognition in eyewitness research, versus the established superiority of recognition over recall in traditional laboratory research.

Other possible contributing factors are subject control and accuracy motivation: Functional and motivational factors operating at the time of remembering may not only be more salient in real life than in laboratory settings (Baker-Ward et al. 1993; Neisser 1988b), but also the assessment methods commonly employed in naturalistic research may allow a greater opportunity for these factors to exert their influence. Consider, for instance, the open-ended, free-narrative methods of eliciting information recommended for questioning witnesses. As discussed earlier, such methods offer subjects much greater control over their memory reporting than is allowed in traditional laboratory research, and this control can have a dramatic effect on memory performance, particularly memory accuracy. Moreover, in naturalistic situations the functional incentives for accuracy are often much stronger than in typical laboratory experiments. In concert, then, these factors could explain the sometimes remarkable recall accuracy observed in naturalistic settings (Hilgard & Loftus 1979; Neisser 1988b). However, these factors can also produce high levels of accuracy under typical laboratory conditions for such a banal task as memorizing a list of unrelated words (Koriat & Goldsmith 1994, Experiment 2).

Is this to say, then, that there are no actual differences between memory performance in real-life and laboratory settings? On the contrary, as just indicated, and in line with Neisser's (1988b) comments, our work leads us to expect marked performance differences between the various social and functional contexts of remembering (especially between naturalistic and functionally "sterile" laboratory contexts) and also helps pinpoint some of the factors contributing to such disparities: differences in memory property, report option, accuracy motivation, monitoring effectiveness, and other aspects of subject control (e.g., control over the grain size of the report). Of course, there are undoubtedly many other important variables that were not addressed in our research. The point is that only by identifying and experimentally investigating such variables can at least some of the differences in memory dynamics between naturalistic and laboratory contexts be demystified and ultimately understood.

Taken as a whole, then, the present article delivers a double message regarding the context-of-inquiry issue. First, as just pointed out, many disparities may be expected between everyday and laboratory findings, but some of these can be clarified by considering the different assessment approaches and functional concerns that are characteristic of each context.

Second, however, the methodological biases prevalent in the study of everyday memory appear to reflect a more fundamental departure from the laboratory tradition, a

departure in terms of the very metaphor of memory espoused. This shift toward the correspondence metaphor is expressed in the preference for complex stimulus materials having an internal structure, in the focus on the many qualitative ways in which memory can change over time, and on the processes underlying these changes, in allowing for the contribution of subject variables and subject control to memory performance, in the study of motivational and functional factors that may affect such contributions, and of course, in the memory property of interest.

In the following, final section, we return to consider the broader ramifications of the distinction between the correspondence and storehouse metaphors, both for the everyday-laboratory controversy and for the study of memory generally.

## 6. Memory metaphors and the real-life/laboratory controversy

In introducing the controversy between proponents of laboratory and naturalistic memory research, we identified its three basic dimensions, referred to as the "what," the "how," and the "where" of memory research: the first of these concerns the content of memory study, that is, the substantive topics deemed worthy of investigation, the second concerns the proper methodology, and the third involves the appropriate context of inquiry. We stressed that although the three dimensions are intercorrelated in the reality of current research practices, they are not logically interdependent, and therefore we sought an implicit common denominator at the metatheoretical level. We then tried to show how part of the cleavage between the traditional laboratory approach and everyday memory research might be captured by the contrast between the storehouse and correspondence metaphors and their respective quantity-oriented and accuracy-oriented approaches to memory. In particular, we focused on the emerging correspondence metaphor and its unique implications for the study and assessment of memory. In concluding this article, then, we first return to examine how this conceptualization helps bind together some of the issues pertaining to the what, where, and how of memory research. We then consider the implications of this analysis for the everyday-laboratory controversy itself.

### 6.1. The guiding role of memory metaphors

Figure 4 sketches a rough scheme depicting some of the interrelationships we assume to exist between *memory metaphors*, *substantive content* (what), *research methodology* (how), and *context of inquiry* (where). Each aspect is represented by a separate node, and the links between the nodes indicate mutual influences and/or constraints. In addition, more specific features of memory research may be seen to fall into one of four quadrants, representing areas of interaction between the neighboring nodes.

This scheme is based on the premise that, as in other areas of scientific inquiry (see, e.g., Arbib & Hesse 1986; Black 1962; Hesse 1966; Hoffman 1980; Kuhn 1979; Leary 1990; Oppenheimer 1956), conceptual metaphors play a primary role in shaping memory research (e.g., Roediger 1980; Tulving 1979; Watkins 1990). Thus we assume that a memory metaphor combines a pretheoretical point of view with the desire to capture the nature of some memory

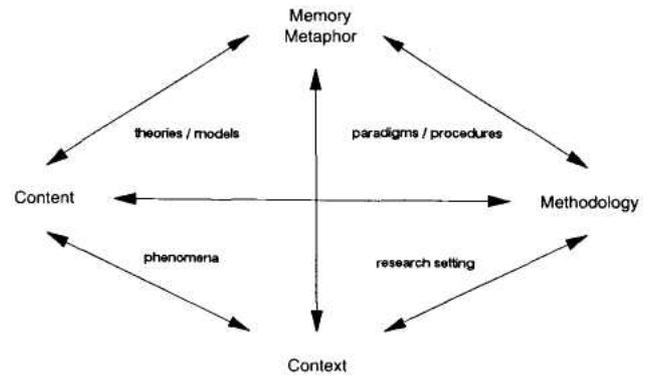


Figure 4. Proposed framework for conceptualizing the relationships between memory metaphors and the "what" (content), "where" (context), and "how" (methods) of memory research.

phenomena. For example, at the metatheoretical level, the storehouse metaphor embodies the empiricist view of the mind as a passive depository of discrete elementary ideas, rather than as an active agent with intentions and goals (see Brewer & Nakamura 1984). At the same time, the metaphor responds to certain basic properties of memory and represents an attempt to understand the intangible mental processes of learning and remembering by drawing an analogy with better-understood aspects of the physical world. Thus the storehouse metaphor embodies the mental analogue of object permanence, the fact that objects deposited in a particular place persist and may later be retrieved. This analogy allows memory to be understood in terms of such familiar physical—spatial notions as depositing, retrieving, losing, searching, displacing, and so forth (Roediger 1980). Note, however, that even a more abstract metaphor, such as the notion of correspondence, may supply its own useful concepts, such as schemata, "goodness of fit," and distortion.

Once adopted, a memory metaphor provides a structured framework within which memory phenomena are analyzed and explained. It can help in abstracting the critical aspects of the phenomena, in defining the substantive questions of interest, and in choosing the methods of investigation. In fact, the metaphor facilitates the development of more specific explanatory theories and models of memory by supplying the basic terms and concepts - the very language of thought in which they are cast. This shared language allows theories based on the same general metaphor to be compared and pitted against each other. When such a common metatheoretical foundation is lacking, however, comparison between theories is much more difficult, if not impossible.

Moreover, together with its associated theories and models, the metaphor will guide the choice of research methodology. For instance, as mentioned earlier, many of the traditional list-learning procedures constitute experimental simulations of the physical process of depositing and recovering elements of information from a memory store. Here, both the experimental techniques and the measures themselves are designed to tap the amount of information retained. By contrast, we have seen that the correspondence metaphor implies very different procedures, namely those that may aid in understanding the factors affecting the congruence between a person's memory of an event and his or her initial perception of that event. Hence, one sees

the preference for output-bound memory measures, complex stimulus materials, subject control, and so forth.

Of course, the picture is more complicated than the discussion so far would indicate, because many of these relationships actually involve mutual constraints. Thus there is an interplay between the metaphor and the phenomena, so that each brings into focus the most compatible aspects of the other. Although the choice of metaphor is certainly constrained by the phenomena, the metaphor itself is selective in drawing attention to those memory phenomena, processes, and variables falling within its "focus of convenience" (Kelly 1955); other topics may be left out.

Furthermore, the research methods and paradigms originally shaped by the metaphor may in turn illuminate and elaborate certain aspects of the metaphor itself. For instance, we showed earlier how the substantial effect of subject control on accuracy-based measures and stimulates a more "active" correspondence conception, leading to a focus on further, substantive issues (e.g., metacognitive contributions) that might otherwise be overlooked. On the negative side, however, when the experimental tools and paradigms are too well established, they may become functionally autonomous (Allport 1961), restricting the range of research topics and constituting a target of study in their own right. In that case, many of the questions asked will be dictated more by the nature of the experimental tools themselves than by any direct relevance to naturally occurring memory phenomena (cf. Conway 1993; Tulving 1979).

With regard to the *context* dimension, the foregoing discussion implies a distinction between two separate aspects: the *context of the phenomena* and the *research setting*. As noted above, an important advantage of adopting a conceptual metaphor is that it can serve as a stepping stone from concrete natural phenomena to their conceptualizations within a more abstract theoretical framework. Thus, although a metaphor may sometimes induce an alienating remoteness of the theories and methods from the original context of the phenomena, it can also encourage a healthy detachment. In fact, such detachment has enabled the great bulk of traditional storehouse-guided research to be conducted in the laboratory, taking advantage of increased experimental control. In the absence of a sufficiently articulate metaphor, it may be more difficult to develop general theories and standard experimental procedures, and the methods of investigation may tend to be more closely tied to the concrete phenomena themselves and their natural contexts.

When viewed from this perspective, the study of everyday memory in natural settings may perhaps be characterized as "phenomena in search of a metaphor," much as the storehouse tradition might be regarded as "a metaphor in search of phenomena." Although the revolt against the storehouse metaphor has enabled everyday memory research to address many new and important memory topics, at the same time such research seems to suffer from the lack of guidance that a well-articulated alternative metaphor might provide. We have argued that the correspondence metaphor explicated in this article can serve to anchor the accuracy-oriented approach implicit in much everyday memory research, yet it is clear that this metaphor has not been articulated or exploited to anywhere near its fullest extent. A further development of the metaphor could facilitate the abstraction of correspondence-oriented ques-

tions from their real-life contexts and perhaps lead to a systematic body of experimental research that is specifically correspondence-based. Instead, there is often a tendency to try to force everyday memory research into the storehouse mold (e.g., quantity-oriented forgetting functions, item-based assessment; see Neisser 1988c; Winograd 1988) or to settle for naturalistic research that is primarily descriptive. Thus, whereas the complex structure and nature of memory phenomena in real-life contexts may be critical in dictating the type of conceptual metaphor appropriate to guide everyday memory research, the research setting, as such, could eventually prove to be the least intrinsic aspect.

## 6.2. Towards a rapprochement: Metaphorical pluralism

According to this analysis, what is the prognosis for the everyday-laboratory controversy? On the one hand, the conceptualization in terms of metaphors implies that the conflict between the two camps may run deeper than is commonly realized. To the extent that it reflects a real metatheoretical breach, it cannot be passed off as an argument about the what, where, and how dimensions, nor can it be reconciled simply by merging empirical findings. In fact, attempts to "be friends" by forcing a consensus regarding surface features of the controversy (cf. Klatzky 1991) may merely serve to obscure an important underlying divergence. We have tried to show that the two different conceptions embodied in the storehouse and correspondence metaphors not only have pervasive implications for how memory research is done, but also reflect fundamentally different ways of thinking about memory. Thus, even if agreement could be reached about the memory phenomena that ought to be studied, the experimental procedures to be used, and the appropriate context of inquiry, the storehouse and correspondence metaphors would still imply different perspectives for looking at the data.<sup>16</sup>

On the other hand, to the extent that the controversy reflects, at least in part, a difference in the underlying memory metaphor, then there really should be no reason for commotion (cf. Tulving 1991). As discussed above, metaphors are cognitive vehicles that help extract the essential aspects of natural phenomena and organize our ideas about them. They reside entirely within the mind of the beholder. Their role is to serve the development of articulate, detailed, and testable theories and models. However, unlike the theories that they breed, metaphors do not make any substantive claims regarding the phenomena that they are purported to capture. Metaphors are neither right nor wrong - their worth depends entirely on their utility. Not only is there no real conflict between metaphors, it is in fact desirable to entertain a variety of metaphors in order to capture the full richness of memory phenomena: Each may contribute to the strength and elegance of the whole.<sup>17</sup>

Thus, although the storehouse metaphor has many advantages that have made it immensely productive in generating memory research and theory, we believe that the correspondence metaphor (perhaps in alliance with the emerging "brain" metaphor; see note 5) has much to offer in capturing those aspects of memory functioning that lie outside the storehouse metaphor's focus of convenience. As argued earlier, exploited to its fullest, the correspondence metaphor could engender a full-fledged psychology of memory correspondence to parallel the quantity-oriented

tradition. This will require, however, the development and systematic use of correspondence measures that can help capture the overall faithfulness of memory under different conditions. It will also require the development of new theoretical frameworks to guide the study of memory correspondence and to uncover the unique variables and processes that may affect such correspondence. This is the type of endeavor that would seem to be mandated by many discussions of everyday memory.

In our opinion, too much has been made of the desire to maintain continuity between the study of everyday memory and the traditional laboratory approach (Klatzky 1991; Winograd 1988). As "metaphorical pluralists," we would advocate otherwise: rather than attempt to bridge the two fundamentally different treatments of memory represented in the traditional and everyday memory camps or perhaps to merge them, we would prefer to see their differences sharpened and cultivated. In that way, each horse will be able to draw the chariot of science as far and as fast as it can.

#### ACKNOWLEDGMENTS

This research was supported by grant no. 032-1106 (1990) from the Israel Foundations Trustees. It was conducted while Morris Goldsmith was supported by a National Science Foundation Graduate Fellowship and completed while Asher Koriat was the Irv Acenberg visiting scientist at the Rotman Research Institute of Baycrest Centre, Toronto, Ontario, Canada. The work was carried out at the Institute of Information Processing and Decision Making, University of Haifa. We are grateful to Robert Lockhart, H. Markowitz, and two anonymous reviewers for their helpful comments on an earlier draft of this article. We also wish to thank Michal Sion for her help in the initial stages of the research and in preparing the manuscript.

Correspondence regarding this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa, Israel.

#### NOTES

1. These labels represent only a rough attempt at categorization and should not obscure the overlap that clearly exists between some of the approaches.

2. Neisser (1984), for instance, declared that "remembering is like problem solving rather than like reproduction. . . . The widely held belief in permanent storage of specific experiences has essentially no basis in fact" (p. 33).

3. This statement was made in criticism of Head's (1920) use of "storehouse" terminology. Bartlett's preceding comments are perhaps also worth reproducing here: "First, Head gives away too much to earlier investigators when he speaks of the sensory cortex as the 'storehouse of past impressions.' All that his experiments show is that certain processes cannot be carried out unless the brain is playing its normal part. . . . One might almost as well say that because nobody who is suffering from a raging toothache could calmly recite 'Oh my love's like a red, red rose,' the teeth are a repository of lyric poetry" (1932, p. 200).

4. A similar approach is implicit in current theoretical explanations of the accuracy and inaccuracy of various "feeling-of-knowing" phenomena (e.g., Begg et al. 1989; Kelley & Lindsay 1993; Koriat 1993; Metcalfe et al. 1993; Reder & Ritter 1992; see Schwartz 1994, for a review).

5. Much as the storehouse conception was bolstered by the computer metaphor, the correspondence conception may perhaps be similarly served by the *brain* metaphor now gaining favor in memory research. An early forerunner of this latter metaphor was Fechner's (1882/1987) suggestion that the psychophysical representation of memories might be conceived in terms of global neural "oscillatory patterns." Other, similar metaphors also seem

to imply a correspondence view, for instance, "resonance" and "holography" memory metaphors (e.g., Gibson 1979; Metcalfe 1990; and see Roediger 1980). Further examples are the "stage-setting" and "re-creation" metaphors of Bransford et al. (1977), in which the role of past experience is to provide "boundary constraints" for memory processes that "are assumed to involve a gradual refinement of one's level of attunement in an attempt to re-create a particular experience. Such processes are therefore assumed to be more like those involved in creating a particular action . . . than they are like a search for some *thing* stored in a particular place (p. 449, emphasis in original).

6. In fact, as opposed to the more gradual, wholistic processes of learning, single-trial "storage" is difficult to implement in connectionist models, though attempts are being made to overcome this lacuna.

7. Indeed, in characterizing the general treatment of subject-controlled processes in traditional memory research, Nelson and Narens (1994) note that memory researchers commonly "attempt to eliminate or reduce their subjects' variations in self-directed processing because (1) such processing on the part of the subject is typically construed mainly as a source of noise . . . , and (2) until recently, there have not been theoretical frameworks within which to systematically explore the subjects' self-directed processing" (p. 9). Of course, there are exceptions, for instance the study of organizational strategies in recall (e.g., Bousfield 1953; Bower 1970; Tulving 1962).

8. Test format may be conceived as a continuum representing the extent to which the possible response alternatives are constrained, ranging from production tests, where subjects produce their answers with little or no intervention (e.g., free narrative, free recall), to selection tests, where one or more stimuli are provided as response alternatives (e.g., lineup or mugshot identification, multiple-choice recognition). Between the two extremes there are procedures that seem to exert intermediate levels of constraint, such as cued recall or directed questioning (e.g., "Was the door open?" or "What model car was it?"). Similarly, particularly in naturalistic situations, report option may also constitute a continuum, representing the extent to which there are explicit or implicit demands to provide an answer to all questions/items (cf. Neisser 1988b).

9. Adding to these considerations, simulation analyses and empirical results also indicate that the *rate* of the tradeoff generally increases as the response criterion level is raised (i.e., under higher accuracy incentives; see Koriat & Goldsmith, in press b).

10. It is important to distinguish between two indices of monitoring ability, calibration, and resolution (see Lichtenstein et al. 1982; Murphy 1973; Nelson 1984; Yaniv et al. 1991; Yates 1982; 1990). Calibration refers to the absolute correspondence between subjective confidence (assessed probability) and the likelihood that an answer is correct, whereas resolution refers to the relative discrimination between correct and incorrect answers afforded by the confidence judgments. Whereas calibration may be important for achieving an absolute level of desired accuracy, it is resolution that is required for the successful screening out of incorrect answers in the service of increasing memory accuracy (see Koriat & Goldsmith, in press b).

11. We use the term "retention" genetically, to mean either the actual retention of information (a storehouse concept) or the ability to recreate (e.g., reconstruct; Neisser 1967) such information when needed (see also "ecphory" in Tulving 1983).

12. Several reports in the literature indicate situations in which monitoring may be rather poor: Cohen (1988), for example, found that although subjects were quite accurate in monitoring the recallability of studied words, their judgments of the recallability of self-performed tasks (SPTs) had no predictive validity whatsoever. Also, Fischhoff et al. (1977), who investigated the calibration of subjective confidence, showed that certain so-called deceptive general-knowledge questions tend to produce an illusion of knowing, engendering an undue confidence in one's incorrect answers. In fact, Koriat (1995) recently found that when subjects

failed to recall the answer to such deceptive items, their feeling-of-knowing (FOK) judgments were either uncorrelated or even negatively correlated with subsequent recognition memory performance. A similar dissociation between confidence and the validity of one's answers has been shown to ensue from postevent misinformation (see Weingardt et al. 1994). Finally, there is evidence indicating that monitoring abilities may be relatively impaired in certain special populations, for example, young children (e.g., Pressley et al. 1987), Korsakoff patients (e.g., Shimamura & Squire 1986), and patients with frontal lobe lesions (e.g., Janowsky et al. 1989).

13. In line with this basic posture, variation in retrieval effort has even been termed "processing bias" (Erdelyi et al. 1989).

14. We should also point out that although the hit and false-alarm rates are sometimes considered to represent quantity and accuracy measures, respectively, these indices are very different than free-report measures: First, unlike the free-report quantity measure, the hit rate can be raised arbitrarily to any desired level. Second, unlike the free-report accuracy measure, which has a straightforward interpretation in terms of the dependability of *volunteered* information, the false-alarm rate has nothing to do with the subject's actual *commitment* to the correctness of his or her answers (see also Koriat & Goldsmith 1994).

15. Interestingly, when signal-detection methods are applied to (forced) *recall* tasks, *d'* reflects monitoring ability alone (i.e., discriminating correct from incorrect answers) rather than retention (see Lockhart & Murdock 1970). Further complications regarding the interpretation of *d'* for various testing methods are discussed in Banks (1970).

16. Consider, for example, the radically different conclusions mentioned earlier about the effects of motivation and recall criterion on memory performance that were reached in our accuracy-oriented studies (Koriat & Goldsmith 1994; in press b) and in prior, quantity-oriented studies (e.g., Erdelyi et al. 1989; Nilsson 1987; Roediger & Payne 1985). Both sets of studies addressed essentially the same phenomena in a common, laboratory context, yet from very different perspectives.

17. Indeed, such an approach seems to have been adopted by the old Jewish sages in their attempt to characterize various different qualities of human memory: "There are four traits among those who sit before the sages: a sponge, a funnel, a strainer, and a sifter. A sponge - because he sponges everything up; a funnel - because he takes in on one side and lets out on the other; a strainer - for he lets out the wine and keeps in the lees; and a sifter - for he lets out the flour and keeps in the finest flour" (Pirke Avot, reprinted and translated from the Hebrew by Neusner, 1984, pp. 162-63).

ence" approach is vulnerable on conceptual grounds and is not applicable to significant portions of empirical research.

I found Koriat & Goldsmith's (K&G's) partitioning of memory research into camps emblemized by the storehouse and correspondence metaphors to be stimulating and challenging. Their views have much merit in diversifying our arsenal of memory measures, but there are difficulties in their construal of the "new wave" of memory research subsumed under the term "correspondence."

K&G argue that the correspondence metaphor treats memory as a perception of the past, and that, in general, "the correspondence conception of memory has much in common with the way we think about *perception*" (sect. 2.2, para. 6). Although I agree that memory and perception have much in common, I do not believe that "correspondence" captures the nature of the mechanism or the processes shared by the two. To understand my reasons, consider how perception is described in several popular current textbooks. According to McBurney and Collings (1977), perception refers to "the processes by which an organism responds to features of the environment with regularities in its behavior" (p. 3); similarly, Goldstein (1989) takes the term "perception" "to refer to all experiences caused by stimulation of the senses" (p. 2). Sekuler and Blake (1994) define perception as "[t]he acquisition and processing of sensory information in order to see, hear, taste, smell, or feel objects in the world" (p. 515). Schiffman (1990) notes that beyond physiological processes, "[p]erceptions are associated with the organization and integration of sensory attributes" (p. 1), as do Coren et al. (1994), who characterize perception simply as the "conscious experience of objects and object relationships" (p. 12).

The definitions vary, to be sure, but conspicuously lacking is any reference to a "correspondence between what we perceive and what is out there," or "the various ways in which they [perceptions] may deviate from reality" (sect. 2.2, para. 6). According to K&G, reality and perception are supposed to correspond where the latter is true, but to fail to correspond where the latter is false. Yet their correspondence view, like naive versions of the correspondence theory of truth in philosophy, fails to take full account of the fact that it is perception alone that substantiates "reality." We lack a privileged route to reality; nor can we perform neutral observations - ones that are not already theory-laden. Therefore, few psychologists today would submit to K&G's criteria for gauging perception: "the extent to which it accords with reality," its "deviation from veridicality," or "the extent it accords with the input" (sect. 2.2, paras. 3 and 5). These were perhaps acceptable during the heyday of positivism, but K&G's own example - multiple percepts of the crime in Kurasa's *Rashomon* - itself undermines a strong version of correspondence.

Perceptions and memories are emergent (de)constructive phenomena, and the discipline of psychophysics is fully compatible with this view. Fechner, in *Elemente der Psychophysik*, sought to develop an exact science of "the functionally dependent relationships" between physical and psychological phenomena. A century later, scientists continue to build upon Fechner's ideas in perception and cognition. Note, however, that "relations" convey more than mere "correspondence." Nor should the "physical" stimulus be confused with "reality." Wilhelm Wundt, Fechner's colleague in Leipzig, distinguished between the natural sciences that are based upon *mediate experience* and psychology, which is based upon *immediate experience*. The specialized instruments used by the former essentially extend the reach of the senses. Thus, both disciplines are based on experience, though on different constructions of experience. The upshot is clear. Physical and psychological variables relate in complex ways; however, because the former is also based on experience, one cannot be absolutely certain what "reality" or "the actual event" is, nor can one know the ultimate meaning of what is being measured.

However, my concerns about the validity of the correspondence metaphor are not solely conceptual. The applicability of this

## Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

### Correspondence conception of memory: A good match is hard to find

Daniel Algom

Department of Psychology, Bar-Han University, Ramat-Gan 52900, Israel.  
f41429@vm.biu.ac.il

**Abstract:** The distinction that Koriat & Goldsmith have drawn between laboratory and naturalistic research is largely valid, but the metaphor they have chosen to characterize the latter may not be optimal. The "correspon-

approach to empirical research may also be limited, as the following example makes clear. Suppose a person recalls loudly the strong low-back pain that spoiled a recent vacation. Can we check the faithfulness of this memory? Because we lack an objective, physical measure of the pain-inducing stimulus, one of the two items needed to test correspondence is missing. Of course, the person could have been asked to assess the referent pain during vacation; then the differences between the two assessments measure the accuracy of the person's memory. That tactic does not really salvage correspondence, however, because we are interested in the person's memory of the original *pain experience*, not in the memory of the perceptual judgment that the person had given at the first assessment. The failure of the correspondence model is not conjectural; it inheres rather in the nature of the tested memory. Does that failure preclude characterizing the memory in question? Not necessarily, as the following study shows.

Algom and Lubel (1994) had separate groups of women estimate the painfulness of selected labor contractions perceptually or from memory. For the latter group, reference contractions were represented by previously associated symbols. For each woman, the pain judgments of individual uterine contractions were related functionally to the biometrically measured magnitude (routine tocographic recordings) of these contractions. Both remembered and perceived painful sensations related to referent pressure by power functions, but the exponents for the memory functions were greater than those for the perceptual functions. Note that despite specification of the endogenous stimulus the correspondence model still does not apply. First and foremost, intrauterine pressure is *not* instantaneous pain; hence, we still lack one item to test for correspondence in *either* perception or memory. Second, both perception and memory are nonlinear functions of pressure intensity, yet we are not justified for that reason to question their faithfulness. We simply do not know. Third, we cannot even say whether perception or memory is more accurate. They simply differ: For one contraction to feel twice as painful as another, its intensity would need to be about 1.64 times as great; for one contraction to be *remembered* as being twice as painful as another, the physical ratio reduces to 1.47. Fourth, pain ratings are idiosyncratic and labile. The same pain may induce different responses in different persons; conversely, different pains may receive the same response with various people. Validity is at issue: Ratings cannot be used safely for comparisons across situations or persons. Finally, all the previous points notwithstanding, pressure and pain reports cannot be compared, for want of a common unit of measurement. Values along the dimensions of pressure (measured in mmHg) and pain (measured by subjective numbers) cannot be classified to those that do and do not correspond.

Nevertheless, memory (and perception) for labor pain could still be described in several illuminating ways. Most important is the specification of the psychophysical scales for perceived and remembered pain referring to a common variable. Intrauterine pressure mapped onto memorial and perceptual representations by the same mathematical functions, yet the parameters governing the functions varied in value. Those functions speak to the dynamic properties characterizing the two system (see Algom & Lubel, 1994, for a discussion of those and other properties). The point to note is that the psychophysical functions do not reside in "the actual event." They, like the functions governing parallel distributed processing, characterize the dynamics of the memory systems under study. Consequently, it seems inappropriate to relate them to the input byway of similarity or correspondence. Of course, the output can be compared to criteria of the investigator's choice (to test for correspondence), but that belongs in the context of evaluation (K&G, sect. 3, para. 16), not in the context of construction.

Let me put my comments in perspective. Koriat & Goldsmith have assembled a broad catalog of methods, theories, and problems in the field of memory. I especially compliment them on the exemplary dissociation of various types of memory measures in their empirical research. The distinction that they draw between

two categories of memory research is based on a forceful argument with which I largely agree. Their article is bound to steer investigators away from a single-minded pursuit of specific memory phenomena by inappropriate means. I believe, however, that their characterization of the "new wave" of memory research is inappropriate. Their "correspondence" metaphor is vulnerable on conceptual grounds and does not apply to a sizeable portion of empirical studies.

## Everyday memory and activity

Richard Alterman

Computer Science Department, Center for Complex Systems, Brandeis University, Waltham, MA 02254. [alterman@cs.brandeis.edu](mailto:alterman@cs.brandeis.edu)

**Abstract:** The target article interprets current psychological research on everyday memory in terms of a correspondence metaphor. This metaphor is based on a reduction of everyday memory to autobiographical and eyewitness memory. This commentary focuses on everyday memory as it functions in activity. Viewed from this perspective, the joining of everyday memory to a correspondence metaphor is problematic. A more natural way to frame the processes of everyday memory is in terms of context, practice, and pragmatics.

This commentary discusses the Koriat & Goldsmith (K&G) target article in light of a critical function of everyday memory: the support of everyday activity. Cognitive theories that are based on the everyday activities of the actor are frequently referred to as "theories about cognition in practice" (see Lave 1988). This commentary draws on an example from the FLOABN<sup>1</sup> project; FLOABN is a computational cognitive model of a pragmatic actor acquiring skill at the usage of mechanical and electronic devices. **Everyday activity.** An example of an everyday activity is: use the department copier to photocopy 50 copies of an exam. Examples of procedural facts that may be known by the individual actor about the department photocopier are:

1. For photocopying a single piece of paper, before pushing the run button, the paper to be copied must be on the glass screen.
2. The power switch is located on the top of the device.
3. The device must be "on" in order to run.

It is the task of the memory system to retain, organize, and deploy the relevant procedural facts in a manner that makes them readily available for current and future problem-solving episodes. Any explanation of the behavior and performance of the memory system must account for the pragmatics of the situation, that is, the task itself, the particulars of the task environment, the dependence of memory on context, and the historic nature of the task.

**Recall strategies.** Recall of relevant task-specific facts from memory depends on context. Part of the context comes from external cues that may be available in the immediate environment. Because much of an individual's expertise is acquired during an activity within a specific task environment, much of the actor's knowledge will essentially be tagged, or indexed, by the external cues available at the time it was learned. Thus, in future problem-solving episodes in that specific task environment, the same external cues may be readily available at a given step in the activity and will trigger the recall of the relevant materials at precisely the right time. Outside the specific task environment, these external cues may not be available.

Another part of the context results from the priming of the immediately preceding actions and events. If the activity is viewed as unfolding in time, then what is recalled determines how the actor acts, which partially determines what new features come into view, which in turn act as tags for the next thing to be remembered. Thus things that are not easily recalled at the outset of the task are readily retrieved as action unfolds.

The correspondence metaphor for everyday memory captures neither the strong dependence of memory performance upon

certain environments, nor the historic and evolutionary nature of memory processes. Memory is not a stand-alone system. Skill at retrieval is coupled with particular tasks and task environments. Memory functions as part of an historic interaction between the internal and external of the actor. What is recalled depends on what is inside the head of the actor, what he was doing, what is outside at the time he tries to recall some fact, what was outside at the time he learned that fact, and what was outside the other times he used that fact.

The pragmatic aspects of everyday memory as it functions in activity also apply to the cases of eyewitness and autobiographical memory. For example, compare the effectiveness of recall at the "scene of the crime" versus "in the courtroom," or the efforts of a witness who rehearses his testimony versus one who does not.

Given the pragmatic nature of everyday memory usage, a correspondence metaphor of everyday memory is unnatural.

**Monitoring.** In the context of everyday memory as it functions in activity, it is more reliable to monitor the course of the activity rather than the accuracy of recall. Some recalled facts may be rejected as irrelevant before they are deployed, but the final arbiter of such decisions is testing it in the real world.

**Accuracy.** Accuracy is not a useful gauge of the performance of the memory system for everyday activities. Suppose the task is to print 50 copies of an exam. A large set of facts will come into play in the course of performing this activity. What does it mean to be accurate? Does it mean remembering correctly the set of facts that are relevant to the task? That is not what is required for the actor to muddle through the situation. A more natural way to measure memory in the course of the photocopying task is to measure the *utility* of the facts that are recalled. Another possible measure is *facility*; that is, how easy was it for the actor to recall the relevant facts?

The storage metaphor and its measure of quantity of recall also seem unnatural. It is not necessary that all the relevant facts be remembered for the actor to accomplish his task. Recalling some of the facts simplifies the task, but, for example, it may be easier to look at the labels on the face of the device to determine where the power switch is rather than attempt to recall that fact.

**The practice of memory.** While the correspondence metaphor may be reasonable for interpreting current psychological practice, the notion of "everyday memory" as it is currently studied in psychology differs from that of the interdisciplinary community of cognitive science. As psychology continues to move outside to study memory in the natural world, it is inevitable that it will move toward more pragmatic theories of memory.

#### NOTES

1. The underlying model of activity in FLOABN is based on adaptive planning (Alterman 1988). Overviews of the project can be found in Alterman et al. (1991; 1995).

## Functional memory requires a quite different value metaphor

Norman H. Anderson

Department of Psychology, University of California, La Mía, CA  
92093-0109. nanderson@ucsd.edu

**Abstract:** The function of memory is to allow past experience to subserve present goal-oriented thought and action. The defining characteristic of goal-oriented approach/avoidance is value. Value lies beyond the reproductive conception of memory that is basic to both metaphors discussed in Koriat & Goldsmith's target article. Functional memory requires a quite different metaphor, for which a grounded theory is available.

The function of memory is to make past experience useful in pursuit of present goals. What are these goals? Not often is the goal to remember some particular object or event. Not often is

reproductive accuracy a primary goal. Instead, the typical goal involves value judgments: approach/avoid, good/bad, and so on. A primary function of memory is to assist in online construction of values in relation to operative goals. Once constructed, moreover, these values may become primary content of memory.

The correspondence metaphor of Koriat & Goldsmith (K&G) seems quite unable to handle this functional conception of memory. Their metaphor is bolted and welded to the concept of accuracy. In contrast, accuracy may not even be defined with respect to values of functional memory. Traditional memory research, similarly, is virtually defined by the accuracy measures of recall and recognition, and so is also inadequate for a functional conception of memory.

As a concrete example, consider your attitude toward everyday memory, say, or toward the K&G correspondence metaphor, or toward this commentary. These concepts represent goals you seek to assess. You place values on diverse aspects of these goals, which involve your personal perspective on memory, outcomes of various experiments, diverse arguments made by proponents and critics, and so forth. These values are integrated into your overall attitude. Although this attitude is a complex knowledge system, much of it can be summarized in a single value along a pro-con dimension.

This attitude is in part constructed online, as you read and reflect on what you read. It itself becomes memory content. Much of the processing underlying your summary evaluation may have only a short-term existence, however, leaving a long-term attitude dominated by the summary evaluation.

This functional approach to memory has been buttressed with an exact methodology. A serial curve of functional memory - which is qualitatively different from the traditional serial curve of recall - can thus be obtained. This uses the methodology of functional measurement, based on a general theory of information integration (IIT). Consider a sequence of informer stimuli, all bearing on some overall value judgment that is given at the end of the sequence. This one final response can, under certain conditions, be fractionated to construct the *serial curve of functional memory*, that is, the curve that specifies the contribution of the informer to that final response at each serial position (see Anderson 1981; 1991; in press).

A critical implication of this functional conception of memory is that the value constructed from given informers may have storage different from the given informers themselves. Hence the serial curve of functional memory may be qualitatively different from the serial curve for memory of these given informer stimuli. This implication has been supported in several independent studies (Anderson 1981, sect. 4.2). In the first of these (see Anderson 1991, Fig. 2) the relevant segment of the serial curve of verbal memory showed pure recency, whereas the serial curve of functional memory showed pure primacy. This outcome demonstrates that traditional approaches based on reproductive memory, including both metaphors discussed by K&G, are essentially inadequate for a functional theory of memory in everyday life.

This functional conception of memory has many of the desirable properties listed by K&G. Foremost, of course, is the functional conception itself. But the present conception is truly functional in everyday life. This has just been illustrated with the value function, for which the accuracy definition imposed by K&G is inadequate.

Similarly, the present functional conception treats the person as an active agent seeking after meaning. But this activity is truly constructive, as just illustrated with value. This constructive view is thus markedly more general than the reconstructive view espoused by Bartlett (1932) and by K&G.

Memory and perception, considered similar by K&G, are also considered similar in IIT. But here again IIT rejects the accuracy criterion. Perception, although related to environmental structure, is essentially different. Perception must be understood in terms of its own laws and organization, which are essentially different from those of the environment. Good progress has been made in integration psychophysics (Anderson 1992; Schlottmann & Anderson 1993; Wilkening & Anderson 1991).

The fundamental mistake of traditional memory theory was to conceptualize memory in terms of reproduction of given stimuli. Despite their well-taken attempts to recognize functional aspects of memory, K&G's emphasis on accuracy shows that they are even more tightly bound to the reproductive conception of memory than proponents of the storehouse metaphor.

Everyday memory is not primarily reproductive memory. The traditional conception of memory as *remembering* has blinded proponents of everyday memory to its essential nature. Everyday memory involves online construction of values and integration thereof. These two operations, valuation and integration, are fundamental to function in general and to memory function in particular. Both operations lie outside the conceptual horizon of either metaphor discussed by Koriat & Goldsmith. Moreover, these integrated values themselves constitute a primary content of memory: This is what memory is for.

## The relation between reproductive and reconstructive processing of memory content

Harry P. Bahrick

Department of Psychology, Ohio Wesleyan University, Delaware, OH 43015.  
hpbahric@cc.owu.edu

**Abstract:** Quantitative losses of memory content imply replicative processing; correspondence losses imply reconstructive processing. Research should focus on the relationship between these processes by obtaining accuracy- and quantity-based indicators of memory within the same framework. This approach will also yield information about the effects of task and individual-difference variables on loss and distortion, as well as the time course of each process.

Koriat & Goldsmith's (K&G's) discussion of the storehouse and correspondence metaphors is an important scholarly contribution. My comments focus on limited issues and should not obscure the high esteem I have for K&G's work.

I do not view the storehouse and correspondence metaphors as mutually exclusive, nor do I believe that the earlier conflict between laboratory and naturalistic memory research continues. Rather, it seems to me that a pluralistic methodology is widely accepted and that the two metaphors are supplementary. Each is suitable for conceptualizing distinct but equally important phenomena of forgetting: the loss and the distortion of memory content, respectively.

Loss of content was the first aspect of forgetting to be explored because that research is methodologically more manageable. It lends itself to laboratory controls and measurements analogous to those used in physics, a discipline that served as a model for the founding of experimental psychology. The subject's task of recalling previously presented material suggests replicative or reproductive memory processes; such processes are adequately conceptualized with the storehouse metaphor.

Reconstructive or inferential memory processing was inferred later by Bartlett (1932), who showed that subjects altered or distorted prose passages systematically so as to bring them into accord with the organization of their existing knowledge. The correspondence metaphor is well suited to conceptualizing reconstructive processing, whereas the storehouse metaphor is not.

Although some memory scholars (e.g., Neisser 1982; 1984) have argued that all memory processing is reconstructive, there is much evidence that both types of processing occur in many tasks and situations. Memory for prose shows evidence of reconstructive processing when gist is recalled and of replicative processing when verbatim quotes are recalled (Dooling & Christiaansen 1977). Tests for retention of a foreign language (Bahrick 1984) show evidence of replicative processing in the recall of individual foreign words prompted by the equivalent English word and of

inferential processing in the comprehension of text that was not previously encountered.

The storehouse metaphor not only helps to conceptualize performance in traditional laboratory tasks of recalling previously presented target items; it also applies to diverse aspects of naturalistic tasks in which there is no systematic loss of correspondence between input and output during the retention interval. Examples include many domains of semantic knowledge acquired in school such as the vocabulary of a foreign language, the names of places, and the facts of arithmetic. Retention of such content diminishes over time but shows no evidence of other systematic changes from input to output.

The value of metaphors lies in their ability to stimulate and focus research, and I see no merit in centering debate or inquiry on the relative validity of the two metaphors. Instead, we should focus research on the conditions that yield each of these effects and, particularly, on the interactions between reproductive and reconstructive memory processing.

I agree with K&G's important observation (sect. 4, para. 1) that the analytic approach allows both accuracy-based and quantity-based measures of memory to be derived and compared within a common framework. Obtaining separate indicators for accuracy and degree of distortion is analogous to the constant and variable error components obtained by the method of average error in psychophysics (Woodworth 1938, p. 396). The constant error is comparable to a measure of distortion; the variable error to a measure of loss. This approach will help clarify a number of important questions regarding the relations between loss and distortion of memory content.

In a recent investigation using this approach (Bahrick et al., in press), college students recalled their high school grades. The results showed no correlation between the degree of distortion (the degree of grade inflation) and accuracy (the number of grades correctly recalled). From this, we concluded that loss of content and distortion of content were sequential rather than interactive. Thus, errors of distortion due to faulty reconstructive processing did not cause the loss of veridical content, but occurred after that content had been lost for other reasons. Extensions of this method will allow us to identify conditions under which loss and distortion of content are independent, sequential processes as opposed to conditions under which distortions cause the losses of veridical content.

The analytic method will also allow us to compare the time course of loss and distortion. This will help us determine the extent to which distortion reflects initial biased perceptual organization of the input and to what extent it is truly a phenomenon of forgetting, that is, a change of input-output correspondence that occurs during the retention interval.

In sum, a research focus on the analytic method described by Koriat & Goldsmith can yield knowledge about individual difference and task variables that bring about loss versus distortion of memory content, about the time course of each of these processes, and about the type and degree of interactions between them. Such findings will greatly enhance our understanding of memory.

### ACKNOWLEDGMENT

Preparation of this manuscript was supported by National Science Foundation Grant BNS-9119800. The author wishes to thank Cathleen O'Toole for many valuable suggestions.

## On correspondence, accuracy, and truth

Ian Maynard Begg

Department of Psychology, McMaster University, Hamilton, Ontario, Canada, L8S 4K1. [begg@mcmaster.ca](mailto:begg@mcmaster.ca)

**Abstract:** Koriat & Goldsmith raise important questions about memory, but there is need for caution: first, if we define accuracy by output measures, there is a danger that a perfectly accurate memory can be nearly

useless. Second, when we focus on correspondence, there is a danger that syntactic correspondence will be mistaken for historical truth.

**Doubts about correspondence.** Koriat and Goldsmith (K&G) provide a thoughtful analysis of tensions underlying the study of human memory in the laboratory and the everyday world. They develop a useful basis for enriching our measurement of memory. Measurement is impoverished if it tells us only how much of an index event is recollected ("quantity"); it is often more useful to know how faithful the overt content of the memory report is to the event ("accuracy").

**Accuracy?** Although K&G have defined their terms operationally, "accuracy" has an aura of virtue and "quantity" sounds like bean counting. They acknowledge that accuracy is under the strategic control of subjects, who can maximize accuracy by recalling fewer items, excluding all but the ones about which they are most confident. Should we adopt this measure of accuracy when grading our students? They could maximize accuracy by reporting only the single piece of information of which they are most certain. Would such a student have a completely accurate memory for the material covered in a cognition course?

**Correspondence?** K&G posit that the traditional view of memory is based on a storehouse metaphor, whereby the mind is "a mental space in which memories are stored and then retrieved by a search process" (Roediger 1980, p. 238). According to the correspondence metaphor, the mind contains representations that refer to past episodes. The representations may deviate from rather than correspond to aspects of the true event, and later reports may deviate still further. Hence, each representation is a perception of the immediate past, and the later reports are perceptions of the more remote past. The constructive nature of representation and the reconstructive nature of delayed report invite literary excursions into postmodernism. But the real question is: What corresponds to what?

Logically, each act of memory includes three informational sets: (A) the objective, external event; (B) the event as received by the witness/participant; and (C) the event as construed at the time of delayed report. With controlled research, we have good information about A, and can draw empirical conclusions about the correspondence between A and either B or C. But once we leave the laboratory, we lose our access to A. We often ask whether, given C, there is any A at all. I have simulated this in research by having subjects rate the probable truth of statements such as "The extended right arm of the Statue of Liberty is 46 feet long." Inside the four walls of the laboratory, the subject's only basis for rating truth is correspondence between what the statement conveys and other contents of mind. Hence the task is to infer the correspondence between C and A on the basis of the correspondence between C and some B that has an unknown correspondence with A.

Statements "ring true" if there is correspondence between C and B, but if that correspondence is increased by factors that are unrelated to A, increments in apparent truth are illusory. We would not buy a second copy of a newspaper to see whether the first one was right, yet statements that repeat information presented earlier in the experiment seem truer and those that contradict earlier information seem false (Bacon 1979). Apparent truth increases even if only the topic of the statement was presented earlier (Begg et al. 1985), and even if the earlier information was expressly presented as irrelevant or false (Begg & Armour 1991). Why? Familiarity breeds truth, although familiarity can arise for reasons that are irrelevant or even inimical to truth. Subjects can exercise control to discount familiarity if the source of the familiarity has been discredited, but the control is not complete (cf. Begg et al., in press); and familiarity can be eliminated if subjects are distracted by having their attention divided when they hear the statements (Begg et al. 1992).

**Truth?** The correspondence metaphor is helpful when we view memory in prospect; given an event, the current representation records it as received, and later references to the record involve

filling in gaps and resolving ambiguities. But when we view memory in retrospect, the decision about whether a current recollection does in fact refer to an actual event cannot be resolved on the basis of correspondence between some aspects of mind and other aspects of mind; truth is neither syntax nor tautology. In the absence of external evidence corroborating the correspondence between the current recollection and the everyday world outside the cognitive system, we must entertain a reasonable doubt about the veridicality of the relationship between the subjective now and the objective then. Common sense says that if it walks like a duck, it's a duck. But all that glitters is not gold. It is dangerous to evaluate truth on the basis of how many of the characteristics of memorial reports the current one has, because the glitter of correspondence may be mistaken for the gold of truth. [See also Libet: "Unconscious cerebral initiative and the role of conscious will on voluntary action" *BBS* 8(4) 1985; Dennett & Kinsbourne: "Time and the observer: *BBS* 15(2) 1992.]

Many examples in the target article refer to courts, but neither quantity nor accuracy has any immediate application unless the index event in question is known to the measurer. The task in court is often to determine whether there actually was an index event that corresponded to the reports of the witness. Courts have a hard job telling whether statements are "true" memories that were merely "refreshed" by suggestive techniques such as hypnosis, drugs, coercive interrogations, and therapy, or "false" memories attributable to those postevent interventions. Can psychologists design a Pinocchio's nose? Courts are increasingly mindful of being "dazzled by bromides served up in a white coat" (*R. v. DEE* 1995) and often scrutinize the scientific basis for psychological expertise. For example, courts have recently refused to admit recovered memories of childhood abuse into evidence because the theory that such events can be repressed and then recovered in therapy lacks a scientific foundation (*State v. Hungerford* 1995), and they have refused to allow a psychologist to express opinions about the credibility of statements based on "statement validity analysis," which is effectively an index of correspondence (*R. v. Jmieff* 1994; *R. v. S.C.H.* 1995).

#### ACKNOWLEDGMENTS

Preparation of this article was supported by Operating Grant OGP0008122 from the Natural Sciences and Engineering Research Council of Canada. I thank Rohan Kristine Robertson for helpful criticisms

## The alternative to the storehouse metaphor

Aaron Ben-Ze'ev

Department of Philosophy, University of Haifa, Haifa 31905, Israel.  
benzeev@research.haifa.ac.il

**Abstract:** Koriat and Goldsmith clearly show the need for an alternative to the storehouse metaphor; however, the alternative metaphor they choose - the correspondence metaphor - is problematic. A more suitable one is the capacity metaphor.

In their excellent target article, Koriat & Goldsmith (K&G) clearly describe the traditional storehouse metaphor for memory as well as an alternative metaphor that seems to explain real-life memory better. Although the distinction between the two models has been made before (for philosophical sources not mentioned in the article, see Ben-Ze'ev 1986; 1993, Ch. 6; Bursen 1978; Reid 1785; Stern 1991), K&G provide important detailed discussions of these models and their implications for psychological research. K&G characterize themselves as metaphorical pluralists, but their sympathy for the alternative metaphor is obvious — and, I believe, rightly so.

My major difficulty with K&G's view concerns the metaphor they choose for describing the alternative to the storehouse

metaphor: the correspondence metaphor. Whereas the storehouse metaphor answers the question, "What is memory?" the correspondence metaphor responds to a different question referring to a certain attribute of memory. As such, the two metaphors are not contradictory, and proponents of the storehouse metaphor can also hold the correspondence metaphor by claiming that memory is a storehouse that fulfills the function of correspondence. Indeed, Locke, the strongest advocate of the storehouse metaphor, spoke about the resemblance of the ideas stored in memory to the things themselves.

Not only can we speak about correspondence in the storehouse metaphor, but such a correspondence is even quite good. In the storehouse model, we store a copy of an event the moment we perceive it; the correspondence between the copy and the event should always be high. Consider the familiar case of people reporting their recollection of a particular event in slightly different versions at various times, although convinced on each occasion that an accurate account of the recollection of the very same event is being given. The storehouse metaphor would have difficulty in accounting for this example, but not because the property of correspondence does not exist in its description of memory; rather, the correspondence between the stored copy and the remembered event remains more or less constant. There is no apparent reason that the event should not be described in precisely the same way on each occasion; after all, the same copy is supposed to be retrieved from storage each time.

The storehouse metaphor of memory is connected with a container approach to the mind that is essentially mechanical: The internal copies are stored in the storage place and then brought to the center of the mental stage. The alternative metaphor to memory is that of capacity, and it is connected with a more dynamic approach, in which the mind is thought to consist of capacities and states. Memory in this view is the capacity of the organism to arrive at states that are similar to its previous states of awareness while preserving a knowledge of their past origin.

In the storehouse metaphor, memory consists of permanently stored entities, which remain more or less stable through the agent's life; in remembering, these entities are reexcited. In the capacity metaphor, the basic elements are capacities and actualized states, and these are retained or preserved but not stored; they are not brought out of storage but are realized or actualized. The capacity to play the piano and the state of being beautiful are retained, but not stored. Storing and retaining are two different forms of keeping something. A storage place is usually conceived as a passive container for holding something. The ideal conditions for such storage are those entailing minimal external influence on the items being stored. We keep an item "on ice," as it were, so that it will remain unchanged until we want to use it again. On the other hand, retention is an active capacity for preserving something. For a capacity or state to be retained in this way, certain activities must be performed. We either use most capacities or lose them.

Among philosophers, Thomas Reid (1785) was the first to present an alternative to the storehouse metaphor. He criticized Locke's description of memory as a storehouse of ideas and defined memory as the capacity to produce a continuance or renewal of a former acquaintance with the thing remembered while believing that some interval has passed. The metaphor used by Reid is that of a habit, which is a land of capacity. Among psychologists, one of the best accounts of the capacity metaphor can be found in Arnold's (1984) excellent book, which K&G fail to mention. Arnold argued that memories cannot be stored like bales. Memories are never "there"; they are potential dispositions to be actualized: "Instead of a storehouse, the more appropriate image would be a constellation of memory dispositions" (1984, pp. 40, 95).

Although the proposed change in the basic metaphor does not diminish the value of K&G's discussion of the two approaches, some modifications should be made. Thus, some of the authors' distinctive attributes for the correspondence metaphor are also part of the storehouse metaphor. These include, for example,

representing a past episode (in the storehouse metaphor, what is retrieved represents past episodes), the ability to represent past events faithfully, and the concern with what is remembered. The correspondence metaphor is also incompatible with the reconstructive approach to memory, considered by K&G to be "the clearest and most productive example" of an alternative to the storehouse approach. A passive retrieval of stored, internal copies is clearly not a reconstructive activity, but the copies can still correspond very well to the events remembered.

Contrary to K&G's contention, memory is not only related to perception in the correspondence metaphor. A quick look at Locke's discussion of the storehouse metaphor clearly shows the affinity of perception for memory. The difference between the two approaches lies in their starting points. In the storehouse metaphor, which appears to be more natural for memory, perception is compared with memory; in the capacity metaphor, memory is compared with perception. In the capacity metaphor, perception is more likely to be described as direct; accordingly, proponents of the direct approach to perception often consider memory to be direct as well (Ben-Ze'ev 1993; Gibson 1979; Reid 1785).

Although I have suggested replacing the correspondence metaphor with the capacity metaphor, K&G's discussion of the two basic approaches to memory remains highly valuable, and most of their claims are compatible with the capacity metaphor.

## Memory, metamemory, and conditional statistics

Robert A. Bjork and Thomas D. Wickens

*Department of Psychology, University of California, Los Angeles, CA 90095-1563. rabjork@psych.ucla.edu;twickens@psych.ucla.edu*

**Abstract:** Koriat & Goldsmith's distinction between encoding processes and metamnemonic decision processes is theoretically and practically important, as is their methodology for separating the two. However, their accuracy measure is a conditional statistic, subject to the unfathomable selection effects that have hindered analogous measures in the past. We also find their arguments concerning basic and applied research mostly beside the point.

Koriat & Goldsmith (K&G) provide an interesting characterization of alternative approaches to the study of memory. They remind us of some important matters concerning how to look at our data in either laboratory or real-world contexts, and they explicate nicely the importance of metamnemonic decision processes as contributors to observed performance. However, we find their characterization of basic and applied research to be disputable and largely irrelevant to their broader issue, and we want to comment on their distinction between free and forced measures of memory and on the limits to their conditional measure of memory accuracy.

In the free-report experiments K&G describe (conventional free recall being the most familiar example), subjects determine what they report of the information they recall or think they may recall. In such experiments, information can be reported correctly or incorrectly, or omitted (RC, RI, and O, respectively). With three categories, performance cannot be completely captured by any single index. Quantity of retention, in the sense used by K&G, is measured by  $P(RC)$ , and conditional accuracy by  $P(RC|R) = P(RC)/P(RC+RI)$ . Neither measure captures how much is reported and how much omitted, or, more importantly, the content of the omitted material. We want to amplify the real-world and theoretical implications of this point.

K&G imply that omissions are not a serious problem in real-world contexts like witness memory. In section 4.2.2.2, for example, they mention a hypothetical witness who correctly recalls that individuals A, B, and C were present at a crime scene but fails to mention individuals D and E. From testifying in a number of cases, we can say with confidence that if other evidence establishes

that D and E were at the scene, the witness will not be considered "accurate" by the jury. The failure of the witness to remember salient aspects of the criminal episode leads juries to lose confidence in what the witness does report.

From both a practical and a theoretical standpoint, it is essential to recognize that K&G's outward-bound accuracy measure is a conditional statistic. It is subject, therefore, to the complex and confusing selection artifacts that have always bedeviled such measures. If different material is reported in different experimental conditions, then such conditional statistics cease to be comparable. For example, in experiments where study conditions (e.g., level of processing) or characteristics of materials (e.g., frequency, concreteness, meaningfulness) are manipulated, each of which has large and known effects on later memory performance, looking at accuracy alone and ignoring what was actually reported can lead to some very strange conclusions. Specifically, subject-by-item selection effects can make the conditional measure  $P(RC|R)$  higher for poor study conditions, more difficult materials, and less-alert subjects.

In the orthographic condition of a levels-of-processing experiment, for example, judging whether the word *pumpkin* is in uppercase letters normally produces abysmal recall of that word, but *if pumpkin* is your special name for your girlfriend, you may be unable to avoid the type of self-referential processing known to produce excellent recall. If you recall only a few such items, then you will show good outward-bound accuracy, but we would not want to conclude that your memory is "accurate" overall. Similarly, a witness who had a good opportunity to view a complex episode and reports many things about it may have lower output-bound accuracy than a witness who had a poor opportunity to observe and reports only a few salient facts. The latter witness should not necessarily be viewed as more credible.

From this perspective, a forced-report procedure is one way to eliminate the selection issue and the problems of conditionalization. This simplification is one of the reasons these procedures are widely adopted, although we do not argue that it makes them superior to free report. In this regard, it is interesting to note that courtroom testimony is seldom free report. Witnesses are rarely permitted to give narrative answers and are often forced to answer questions. As the American legal system has evolved, it is deemed important that both the prosecution and the defense get answers to everything they ask, yielding something close to a quantity measure.

We also take issue with K&G's version of the storehouse metaphor. For expository purposes, they describe a simplified version of this metaphor, but their version distorts its current form, which we see as closer to a library metaphor. An overworked librarian (some kind of central executive in memory) has a complex and individualistic filing system based on experience. Additions to the library are filed in terms of where they fit, as determined by their relationships to what is already shelved. If the librarian is too taxed by other demands when new materials are received, or if the new materials defy ready classification in terms of existing materials, then the new materials do not get shelved in permanent storage at all. At the time books are requested, the librarian uses the cues provided (title, topic, author, publisher, etc.) to try to find a given volume, but those cues may underspecify the target volume or may define a competing alternative better than the desired volume. The organization of the library changes over time, induced by the continuing addition and retrieval of materials, so that volumes that are retrievable at one time may not be retrievable at another. This version of the "storehouse" metaphor captures more of the dynamic and content-specific character of memory than does the less realistic version cited by K&G.

Finally, we find the laboratory-real-world aspect of the metaphor somewhat irrelevant. To make the enormous advances they did, physicists, chemists, and biologists worked with simple materials - smooth balls and inclined planes, not jagged rocks falling down rugged hillsides. Why then should we study memory only in the full social, emotional, and semantic complexity of everyday

life? Real-world situations suggest general principles that are worked out in simplified domains. Controlled research on memory with simple materials has not misled us about basic functional relationships. The phenomena and principles generated by laboratory research - such as the effects of spacing, the number and nature of study trials, the length and content of retention intervals, the differences between recall and recognition or between concrete and abstract materials, the effects of meaningfulness, the differences between discrimination and criterion  $\beta/d'$  and  $\beta$ ), and the types of interference and transfer - have routinely been replicated in real-world memory settings, often in enlarged form.

## The correspondence metaphor: Prescriptive or descriptive?

Darryl Bruce

Department of Psychology, Saint Mary's University, Halifax, Nova Scotia, Canada B3H 3C3. [dbruce@husky1.stmarys.ca](mailto:dbruce@husky1.stmarys.ca)

**Abstract:** Koriat & Goldsmith's abstract correspondence metaphor is unlikely to prove useful to memory science. It aims to motivate and inform the investigation of everyday memory, but that movement has prospered without it. The irrelevance of its competitor - the more concrete storehouse metaphor - as a guiding force in memory research presages a similar fate for the correspondence perspective.

Koriat & Goldsmith's (K&G's) target article is scientific scholarship at its best, and there is much to admire about their correspondence perspective on memory. Nevertheless, I confess to reservations about what it is likely to accomplish. Metaphors in science serve mainly two purposes: They clarify and they generate research. How does the correspondence metaphor stack up on these two counts?

The principal issue that it attempts to sort out is the difference between laboratory and naturalistic memory research. Although K&G state that there are three aspects to the controversy - what memory phenomena are to be investigated, how, and where - only the first is germane to the conflict. As Neisser (1978) saw it in his well-known critique of memory research, the topics investigated since the time of Ebbinghaus had been prompted far too much by the phenomena, paradigms, and theories of the laboratory and far too little by the problems, puzzles, and demands on memory of everyday life. Recognizing the natural contexts of memory would inevitably lead to a broader array of questions and a wider recognition of significant variables. That was the issue, nothing more, nothing less.

Now come K&G, with their contrast between the storehouse and correspondence metaphors, suggesting that it maps onto the distinction between laboratory and everyday memory research and that the conflict between them runs deeper than is generally realized. But to apply the contrast between the storehouse and correspondence metaphors to the dispute between advocates of the laboratory and everyday memory approaches, though it may make some sense of the debate, is to apply it to a controversy that is over. One need only scan the prevailing state of memory research to see that it is a far broader mix of problems, methods, variables, and theoretical orientations than the comparatively narrow study of memory that marked the first one hundred or so years of our science. In brief, what K&G have provided us with is a post-mortem. Intriguing as that may be, I would have found it more compelling had the correspondence metaphor clarified a live issue.

The second question is whether the correspondence idea will be an engine for memory research. Only time will tell, but I am dubious. Why? Because directions have been taken in memory research - some recent and some not so recent - that fly in the face of the storehouse metaphor; yet the moves have been made

by those who seem party to that metaphor and who are probably unaware of the correspondence alternative. Here are but three examples.

One is a recent investigation of memory accuracy by Roediger and McDermott (1995). Based on an earlier study by Deese (1959), the research stands squarely in the laboratory tradition and retains many of the accoutrements of the storehouse metaphor — lists of words, free recall, and yes/no recognition tests. Yet the focus is on errors of memory - the recall of words that were not presented - and as K&G note, a concern for such errors indicates an emphasis on accuracy or correspondence. Was this work stimulated by the correspondence metaphor or disenchantment with the storehouse metaphor? I think neither. More likely, it arose out of a growing interest in memory distortions per se, heightened no doubt by societal concern about false memories.

A second example is research on the phenomenal characteristics of remembering, that is, the subjective states of awareness that accompany memory reports. Gardiner and Java (1993) requested that subjects make "know" or "remember" judgments to accompany conventional recognition decisions. Know judgments reflect a feeling of familiarity; remember judgments, a recollection of having experienced an item when it was presented. Such judgments are readily dissociated from measures of quantity. Whether they are measures of accuracy is perhaps debatable, but they appear to possess some of the attributes of the correspondence metaphor: They refer to the quality of memory and are outward bound. The inspiration for such an approach? Gardiner acknowledges Tulving (1985), a well-known devotee of memory systems. So the storehouse metaphor hardly seems to have been inhibiting in this instance.

My third example is research on the temporal dimension of memory, something that K&G indicate is foreign to the storehouse perspective but compatible with the correspondence metaphor. In citing a number of everyday memory studies of the problem, they curiously overlook traditional laboratory research on the topic (most of it conducted earlier than the studies they do mention) by those who must be deemed to have held to a storehouse conception of memory (e.g., Underwood 1977). Again, how constraining or guiding has that metaphor been?

The point of these examples is that memory research moves in directions that are independent of abstract background metaphors. The storehouse idea, arguably one that dominates the laboratory tradition of memory research, has not prevented individuals of that stripe from attending to aspects of memory that imply a correspondence perspective. In that respect, the storehouse metaphor has been irrelevant. Why should we expect any more of the correspondence metaphor? At best, it will be descriptive of memory research, not prescriptive.

I am pessimistic about the effectiveness of the correspondence idea for another reason. Consider the levels-of-processing framework ( Craik & Lockhart 1972). It was put forth as a replacement for the idea of separate memory stores. One subscribed to a levels-of-processing framework or to a memory-stores framework but not both. They were incompatible perspectives, or at least so it seemed. In a way, I wish that Koriat & Goldsmith had adopted that strategy. They could certainly have done so. But theirs is a more tempered, rational, and moderate approach - pluralistic, as they call it. They carve out a portion of the terrain for the correspondence metaphor and another for the storehouse metaphor and suggest that each should be explored to yield a richer picture of memory. Reasonable as that call is, it is unlikely to lift the audience to its feet. Put differently, would we still be quoting Neisser if he had said "if X is an interesting or socially significant aspect of memory, then psychologists should study X more often." Unfortunately, science is not often moved by sensible and modest appeals. It resonates to more extreme statements, maddening though that may be.

## What do memories correspond to?

Martin A. Conway

Department of Psychology, University of Bristol, Bristol BS8 1TN, England.  
m.a.conway@bristol.ac.uk

**Abstract:** Neither the storehouse nor the correspondence metaphor is an appropriate conceptual framework for memory research. Instead a meaning-based account of human memory is required. The correspondence metaphor is an advance over previous suggestions but entails an oversimple view of "accuracy." Freud's account of memory may provide a more fruitful approach to memory and meaning.

One hundred and ten years after the publication of Ebbinghaus's *Memory: A contribution to experimental psychology* (1895/1964), it is disappointing to note the very few theoretical advances that have been made in understanding human memory. Set against this is a century's worth of empirical findings. Indeed, the exponential growth in laboratory experiments now presents us with an array of data so diverse and complex that no researcher could ever hope to make much sense of it all. It is a curious position in which to find ourselves: no theory but a hundred years of data. Proof, if proof were needed, of Wittgenstein's (1980) assertion that, in psychology, theory and method pass each other by. In this context Koriat & Goldsmith's (K&G's) attempt to discern a theoretical basis to current practices in memory research is surely welcome. Their proposal, however, that laboratory research assumes a storehouse metaphor whereas everyday memory research assumes a correspondence metaphor (sects. 2.1 and 2.2), although an advance over previous treatments of these two types of research, nevertheless avoids consideration of some crucial issues.

A central concern of cognitive psychology is the problem of representation; indeed, this emphasis on representation is one of the important ways in which cognitive psychology differs from its predecessor, behaviourism. Yet much laboratory-based memory research bears a striking similarity to research in the behaviourist tradition. Researchers in both traditions are frequently preoccupied with proscribed and specific data sets and the procedures whereby the data were collected, at the expense of theory development. Instead of theory, certain fundamental assumptions are made that facilitate the generation of accounts of specific findings: Such accounts are nearly always limited to the data they "explain" and even here they have only a brief life. One of the fundamental assumptions aiding the process of data explanation in memory research is the notion of a memory *trace*. This assumption neatly and fairly completely finesses any consideration of the representation of knowledge in memory. It seems to me that the concept of a memory *trace* has the same ontological status in memory research as the concept of a *response* in behaviourism - both terms preempt theorizing about underlying representations and processes. K&G's argument that laboratory research is essentially concerned with counting traces (as these manifest themselves in the recall and recognition of previously presented items) is undeniably correct. Their claim that such research assumes, perhaps implicitly, a storehouse metaphor seems equally on the button: Where else could memory traces be kept unless in a store?

Does everyday memory research, which according to K&G assumes a different metaphor - a correspondence metaphor - fare any better with the problem of representation? My view is that it has at least made a start. But the way forward is difficult and there are many enticing byways, one of which, illustrated in section 4 of the target article, lies in investigating observable attributes of an event.

Implicit in this is the belief that what is observable is what the rememberer actually encodes. This leads to oversimple conceptualizations of accuracy and correspondence. There are, after all, many ways in which a memory can be accurate. For example, a memory may be perfectly accurate in that it corresponds directly and fully to some knowledge structure created during a particular experience. It does not follow, however, that that knowledge structure is a veridical, complete, or accurate record of the event.

Furthermore, and troublesome for K&G's argument, some of the best current research and theory on memory accuracy has been conducted *in the laboratory* by Johnson and her colleagues (cf. Johnson et al. 1993). Moreover, Johnson et al. place a heavy emphasis on representation in their theorizing and demonstrate in their empirical work that what is encoded is a synergy of externally and internally generated knowledge.

More generally, in the study of autobiographical memory, for example, it has proved useful to consider memories as transitory mental representations drawn from many different sources of knowledge, (Conway 1990; 1992; 1996; Conway & Rubin 1993). On this view, memories are *interpretations* rather than literal records of the past: interpretations that have meaning - personal meaning - *in the present* for the individual rememberer and are intrinsically bound up with self. Taking this view, issues relating to accuracy (as that term is used by K&G) are less pressing, and precedence is given to questions concerning the nature and meaning of the representations. Neither storehouse nor correspondence metaphors are appropriate for this approach to memory, according to which memories correspond to meanings rather than to externally verifiable attributes of events.

It is perhaps ironic that at the same time that Ebbinghaus produced his influential book, other equally important works on memory were also appearing. I am thinking here of Freud's early work on memory (in Breuer & Freud 1895/1955) and, subsequently, his prescient account of memory in Chapter 7 of *The interpretation of dreams* (1900/1953). The former work encompasses a correspondence metaphor and the latter goes much further than this. As Freud realized, a theory of human memory must include an account of the representation of knowledge as well as an account of the processes that modulate knowledge access. In retrospect it is regrettable that Freud's sophisticated thinking on memory had so little impact on memory research: a union of Ebbinghaus and Freud might well have given us a hundred years of theory *plus* a hundred years of data and taken us beyond both storehouse and correspondence metaphors towards a meaning-based account of human memory.

## The real-life/laboratory controversy as viewed from the cognitive neurobiology of animal learning and memory

Howard Eichenbaum

Center for Behavioral Neuroscience, SUNY Stony Brook, Stony Brook, NY 11794-2575. [howard.eichenbaum@sunysb.edu](mailto:howard.eichenbaum@sunysb.edu)

**Abstract:** Parallel to Koriat & Goldsmith's accounting of human memory, there are two distinct approaches in animal learning. Behaviorist approaches focus on quantitative aspects of conditioned response probability, whereas cognitive and ethological approaches focus on qualitative aspects of how memory is used in real life. Moreover, in animal research these distinguishable measures of memory are dissociated in experimental

Koriat & Goldsmiths (K&G's) distinction between "traditional laboratory" and "everyday" human memory has an interesting parallel in approaches to understanding memory in animals. K&G contrast laboratory studies, focused on the quantity of memory stored, and "natural" memory research, focused on the quality of the memory expressed. There are very similar issues at the heart of a longstanding controversy in animal memory research in which traditional behaviorist or "stimulus-response" (S-R) theory is contrasted with "cognitive" and "ethological" approaches to memory. Behaviorist research has primarily pursued quantitative measures, focusing on the incidence of learned responses to arbitrarily assigned conditioning stimuli. Running against this trend (and getting less attention historically), has been the "cognitive" ap-

proach, (e.g., Tolman), as well as recent ideas from ethologists who characterize learning as modifications of instinctive behavior (Gould & Marler 1987). The cognitive and ethological approaches differ on several points, but are both similar to "everyday" human memory research in that they focus on the qualities of memory expression rather than the incidence of "correct" responses. Given these superficial similarities with the real-life/laboratory controversy in human memory, does the literature from cognitive neurobiology have anything more fundamental to say about K&G's account?

The behaviorist or S-R approach, like laboratory research on human memory, has its main advantage in tight experimental control over stimulus dimensions and response measures. It can closely track the acquisition of the conditioned response to the conditioning stimulus, and it can manipulate a full range of imaginable dimensions to account for the rate of learning. However, this strategy can be quite limited when applied to some of the most interesting experimental questions. For example, we almost lost the hippocampus to animal memory research because of behaviorist theory (Cohen & Eichenbaum 1993). Early studies on human amnesia showed that the hippocampus is critical to memory. However, the initial attempts to model amnesia in animals, based strongly on S-R approaches, failed to reproduce the "global" memory impairment observed in humans with hippocampal dysfunction. Indeed these studies almost convinced us that the hippocampus had nothing to do with memory per se, at least in animals. Fortunately, O'Keefe and Nadel (1978), adopting Tolman's notions about cognitive mapping, were able to associate the most interesting kind of memory with hippocampal function in rats. In particular, they found that when the hippocampus is removed rats cannot learn spatial maps of the environment, although they are fully intact in S-R learning that mediates incremental acquisition of approach responses to specific rewarding stimuli. [See multiple book review of O'Keefe & Nadel's *The Hippocampus as a Cognitive Map*, *BBS* 2(4) 1979.]

O'Keefe and Nadel saw hippocampal function as dedicated to cognitive maps for literal space. This turned out to be too limited a view: recent studies have shown that the hippocampus is critical to organized memory representations for all sorts of material (Cohen & Eichenbaum 1993; Eichenbaum et al. 1994). Moreover, the evidence for a general hippocampal memory function comes from observation on the "real life" of rats, and I will give two examples. One involves a natural situation in which rats learn from conspecifics what foods are good to eat. The critical memory involves an association between a novel food odor and an odorous constituent of rats' breath. Rats with hippocampal damage can initially acquire the association, similar to intact short-term memory in human amnesics, but they fail to demonstrate long-term expression of this association when required to use it in guiding food selection (Bunsey & Eichenbaum, *in press a*).

In another recent experiment, we (Bunsey & Eichenbaum, *in press b*) found that rats can also learn olfactory "paired associates" in a more formal framework based on their natural foraging behavior. In probe tests following paired associate acquisition, normal rats demonstrate remarkable capacities for transitive inference between items that share a common associate and for identifying symmetrical relations between associated items presented in the reverse of training order. Rats with damage to the hippocampus can acquire the initial paired associations at the normal rate but they fail to demonstrate either form of flexible expression of these memories. Thus their ability to increment appropriate choice responses to paired stimuli did not depend on hippocampal function, but the hippocampus was required for access to these representations beyond mere repetition of the learning event.

Expanding on these findings, it is now clear that multiple memory systems encode different aspects of a single experience. For example, in a task where rats search for food on a "radial maze" (shaped like the hub and spokes of a wheel), three different forms of the relevant knowledge are developed (McDonald

& White 1995). One, mediated by the hippocampus, is especially attuned to the navigational routes taken. The amygdala mediates a representation of biases toward and against rewarding and nonrewarding loci respectively. And the dorsal striatum mediates an S-R encoding, that is, increases the likelihood of approaching rewarded loci. Another example: multiple systems also mediate different aspects of conditioned stress responses to fearful stimuli (Phillips & LeDoux-1991). The hippocampus mediates a representation of the background context where the fear-producing stimulus occurred, whereas the amygdala mediates acquisition of fear responses both to the context and to particular cues that were in immediate attention at the onset of the frightening stimulus.

Framing these observations on parallel memory systems in the context of K&G's account, it seems the hippocampus is especially important for the unique qualities of "everyday" memory, whereas other systems mediate the adoption of biases about stimuli and increments in conditioned responses in typical laboratory tasks. In other words, cognitive neurobiology goes one step further than do K&G. Experimental amnesia may provide a basis for directly dissociating "quantitative" properties of memory, measured primarily by changes in response-probabilities, from "qualitative" properties, revealed in the flexible expression of memory.

This may be an overinterpretation of K&G's account. However, it offers some pause for thought, and includes a specific suggestion. The connection between the human and animal literatures might be drawn closer by considering the literature on "implicit" versus "explicit" memory in humans; these are broadly viewed as the equivalents of S-R and cognitive memory, respectively, in animals (Cohen & Eichenbaum 1993). My prediction is that implicit memory will turn out to be more sensitive to "quantitative" variables, whereas explicit memory will better capture the "qualitative" dimensions of everyday memory.

## Implications of output-bound measures for laboratory and field research in memory

Ronald P. Fisher

Department of Psychology, Florida International University, North Miami, FL 33181. fisherr@servas.fiu.edu

**Abstract:** Everyday memory tasks often require that researchers focus on output-bound measures of memory. As a result, nonmemorial processes (e.g., report option and grain size) may influence recall accuracy. These nonmemorial processes, usually eliminated by laboratory researchers, have the potential to explain some anomalous results and may even be useful to enhance everyday recollection.

Koriat & Goldsmith's (K&G's) novel analysis, that the contrast between laboratory and everyday memory research reflects a more fundamental distinction between input-bound measures of quantity (storehouse metaphor) and output-bound measures of accuracy (correspondence metaphor), assuredly elevates the laboratory-everyday memory debate to a higher plane than we have seen in recent years. The argument presented here (sect. 2.2, paras. 8 and 9) and elsewhere (Koriat & Goldsmith 1994) is that there is a nonarbitrary link between research in everyday memory and the use of an accuracy-oriented approach, that is, there is something inherently compatible about the marriage between everyday memory and the correspondence metaphor. I suspect that, in addition to any functional relation, there is also a computational reason for the frequent use of output-bound measures; specifically, input-bound measures cannot be calculated meaningfully for some everyday memory tasks.

In a typical eyewitness memory study, experimental subjects observe an action event (e.g., bank robbery) and then attempt to describe it. We cannot enumerate all the details within the event,

as there are, literally, an infinite number of propositions describing the event. Therefore, we cannot calculate what percentage of the event's details have been captured in the response (input-bound measure). Rather, we are relegated by default to measuring only the witness' response (output-bound measure). Thus, we focus on the witness' output, and not on the input, in large part because we can specify the entire contents of the output but we cannot do so for the input.

Because of this natural focus on the witness' response (output-bound measures), the everyday memory researcher forfeits considerable experimental control over the object being measured. Response accuracy, we are told, is extremely sensitive to the witness' choice of whether or not to volunteer a response (report option) and how precisely the response should be reported (grain size). Three interesting consequences emerge from this interdependence between accuracy, report option, and grain size.

First, it suggests that in order to meaningfully compare response accuracy across two experiments, one needs to ensure that the witnesses in both experiments were similar on the dimensions of report option and grain size. Practically, it is difficult to imagine how, in any real eyewitness setting, one could measure these properties. This implies that comparisons of accuracy rates across real eyewitness settings are tenuous at best. This is particularly important in some legal cases, where jurors may be convinced by expert witnesses that because laboratory witnesses are often inaccurate it can be concluded that real witnesses will be equally inaccurate (Fisher 1995).

Second, the interdependence between measures may explain some experimental anomalies that are created by looking at only one measure in isolation. In my lab, for instance, we were perplexed by the odd finding that recall accuracy (proportion of responses that are correct) was not affected by retention interval. In three separate eyewitness studies, recall accuracy was virtually identical after short delays (one or two days) and long delays (several weeks). We have just rescored the data and found that the responses made after long delays were less precise (coarser grained), although equally correct, compared to those provided after short delays.

Third, report option and grain size are decisions that are made *after* witnesses have brought into conscious awareness their recollection of the event ("*ecphory*"; Tulving 1983). Hence, report option and grain size are principles of communication and not memory per se. In laboratory studies of memory, we normally try to minimize the effects of nonmemorial factors, for example, by limiting grain size to the experimenter-determined choices or by conducting a signal detection analysis to examine *d'* without the contaminating effect of beta. In everyday memory research, however, we often do not have the luxury of eliminating or controlling these nonmemorial factors, and so they become an integral part of the eyewitness recollection process. In fact, I suspect that these postephoric processes are probably easier to manipulate, because they are implemented with greater conscious control than are the more automatic processes of *ecphory* (Klatzky 1984). We may accordingly find that the best way to enhance eyewitness recollection in the real world is by applying principles of communication or other nonepchoric processes rather than principles of memory. This stands in stark contrast with the laboratory approach of trying to minimize the contribution of nonepchoric processes when studying recollection.

It seems safe to conclude that K&G's lucid development of the correspondence metaphor as an alternative to the traditional storehouse metaphor will create several new insights into the working of memory. It is probably too early to predict where these insights will lead, but one thing we can be sure of is that the intellectual ride will be exciting.

## Driving and dish-washing: Failure of the correspondence metaphor for memory

Keith S. Karn<sup>a</sup> and Gregory J. Zelinsky<sup>b</sup>

<sup>a</sup> Center for Visual Science, University of Rochester, Rochester, NY 14627-0270; <sup>b</sup> The Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801.  
keith@cvs.rochester.edu; greg@cvs.rochester.edu

**Abstract:** Koriat & Goldsmith restrict their definition of memory to "being about some past event," which causes them to ignore the most common use of memory: everyday visual-motor tasks. New techniques make it possible to study memory in the context of these natural tasks with which memory is so tightly coupled. Memory can be more fully understood in the context of these actions.

Although Koriat & Goldsmith (K&G) propose an interesting, and probably useful, *correspondence* metaphor for the study of memory, their definition of memory as "being about some past event" or "episode" seems overly restrictive. Instead of being dedicated to providing eyewitness testimony, recalling autobiographical events, or recognizing previously viewed images, memory is most often used, without conscious awareness, in natural tasks such as driving, walking, grasping, speaking, and problem solving. The correspondence view fails to describe this repertoire of behaviors because it, like the storehouse metaphor, maintains a static conception of memory as a retrieval device without recognizing the active role that memory plays in goal-directed behavior.

For example, the task of driving along a familiar road may seem to be almost automatic. During this task, however, we rely on memory for complex motor sequences (e.g., hand movements to shift gears or eye movements to check the rear-view mirror), for decisions as to which direction to turn, and for direct interactions with the environment (e.g., stopping for lights or pedestrians). Nevertheless, we are able to reach our destination, often with minimal awareness of the events that transpired during the process. It is this reliance on memory, of which the person may not even be aware, that K&G (and most others) fail to consider.

Although you could stretch the correspondence metaphor to describe almost any situation requiring memory, there are some fundamental differences between memory use in performing visual-motor tasks like those described here and in recalling past events. The memory use we refer to requires frequent, rapid access and then equally rapid disposal of information (with a time scale of a few seconds) and is inextricably linked to our moment-by-moment behavior.

Our daily activities are filled with examples of this sort of memory use. Washing dishes is another task performed almost automatically, without consciously relying on memory. This task requires that visual information be acquired from the environment (e.g., type of dish, location of soap, etc.) and also that appropriate procedural routines be recalled from longer term memory (e.g., delicate handling for a wine glass, forceful abrasive action for an iron skillet, etc.). Both types of information must be used for a few seconds, after which the information must be purged and new information acquired and stored so as not to misapply the pot-scrubbing technique to a fragile glass.

This type of memory use has often been ignored, perhaps because it cannot be assessed by standard techniques. Recently, however, Ballard, Hayhoe, and their colleagues (Ballard et al. 1992; 1994a; Hayhoe et al. 1994; Pelz et al. 1993) have developed new laboratory techniques for studying this real-time memory use in natural tasks without the perturbation caused by recall or recognition probes. In this technique, eye, head, and hand movements are monitored during a block-copying task. The subject constructs a copy of a visible *model* pattern of randomly arranged, colored blocks in a *workspace* by moving blocks from a *resource* area (see Fig. 1).

Frequent refixations of the model during the task (typically twice for each block moved) indicate that memory plays a critical role in maintaining color and location information, but for very

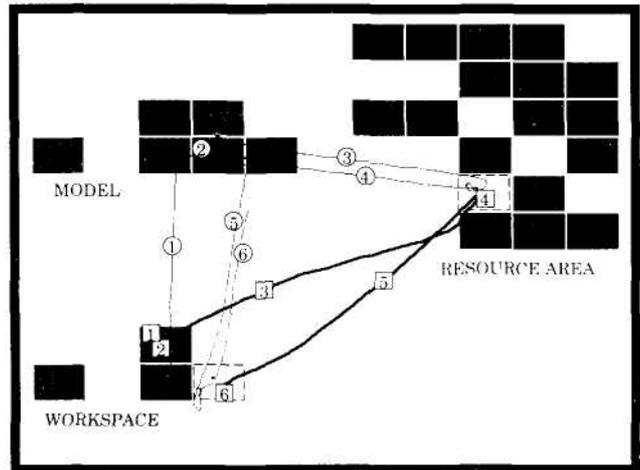


Figure 1 (Karn and Zelinsky). A typical pattern of eye (thin line) and hand (thick line) traces during the movement of a single block in the block copying task. Saccadic eye movements are numbered, as is the corresponding hand position at the midpoint of the saccade. Two fixations in the model area during the movement of a single block are typical.

brief periods (Ballard et al. 1994a; 1994b). Furthermore, changing the color of model blocks during a saccade reveals dramatic differences in fixation durations depending on the particular phase of the task (i.e., block pickup vs. block placement) in which the change occurs (Bensinger et al. 1995). From this we can infer that the accuracy of color information held in memory varies during different phases of the task. When previously placed workspace blocks are removed during a saccade such that they are not visible during the fixation immediately preceding a targeting saccade into the workspace, the precision of the targeting eye movement degrades only slightly, even though the saccade is launched toward an invisible target. This suggests that rather precise spatial information is held in memory just prior to block placement (Karn 1995). These data support the notion of a close tie between working memory and the information required at each moment during a task.

In summary, we have argued that the necessity to recall actual episodes is relatively rare, and in tying their definition of memory to such tasks, K&G fail to address the most common uses of memory. Each of us relies on memory thousands of times per day by extracting and storing small pieces of information from the environment or recalling motor routines for rapid use and disposal. This more common type of memory use has recently been studied in the laboratory by monitoring eye, head, and hand movements with minimal disruption to the natural tasks with which memory is so tightly coupled. We conclude that memory, like perception, can be more fully understood in the context of action.

## The storehouse/correspondence partition in memory research: Promises and perils

Arie W. Kruglanski

Department of Psychology, University of Maryland, College Park, MD 20742

**Abstract:** The novel correspondence metaphor outlined by Koriat & Goldsmith offers important advantages for studying critical issues of memory-accuracy. It also fits well with the current emphasis on the reconstructive nature of memory and on the role of cognitive, metacognitive, and motivational factors in memory performance. These positive features notwithstanding, the storehouse/correspondence framework

faces potential perils having to do with its implied linkage to the laboratory/real-life controversy and its proposal of studying correspondence issues in isolation from memory phenomena captured by the storehouse paradigm.

The novel correspondence metaphor outlined by Koriat & Goldsmith (K&G) offers several compelling advantages. It focuses research attention on issues of accuracy (of paramount relevance to everyday memory concerns - e.g., eyewitness testimony); it emphasizes postevent cognitive, metacognitive, and motivational influences on the remembering process; and it highlights the active role of the rememberer, implying an affinity between memory research and work on perception and judgment.

The emphasis on accuracy affords, in addition, the development of useful new methodologies for memory research. These are exemplified by Koriat and Goldsmith's (1994) distinction between memory property, report option, and test format and the possibility of importing the novel data analytic techniques developed in the realm of social perception to memory research (Kenny 1991; Kruglanski 1989).

Those positive aspects of the correspondence metaphor notwithstanding, K&G's analysis raises several problematic issues. One of these pertains to the laboratory versus real-life controversy in memory research and the risk that the correspondence metaphor will be intimately linked with that particular dichotomy. Such a linkage seems implicit in the authors' comment that "everyday memory research has shifted toward a *correspondence* metaphor" (abstract), that "inherent differences in the dynamics of remembering between everyday and laboratory contexts . . . may limit the generalizability of results across the two contexts" (sect. 5.4, para. 3), or that "part of the cleavage between the traditional laboratory approach and everyday memory research might be captured by the contrast between the storehouse and correspondence metaphors" (sect. 6, para. 1). All this, despite their emphasis that the "what," "how," and the "where" of memory research, though "intercorrelated in the reality of current research practices," "are not logically interdependent" (sect. 6, para. 1).

The latter position implies that it should be possible to study accuracy issues in the lab (as Koriat and Goldsmith [1994], Ross [1989], and many others have done), incorporate functionally meaningful research material into experimental research, or study storehouse-based issues of encoding, storage, and retrieval in realistic settings. If so, neither metaphor should be doomed to the artificial study of functional irrelevancies and neither should have the monopoly on what is useful and significant about human memory. More specifically, the rift between "everyday" and "laboratory" approaches seem to confuse what may have been *typical* with what is *essential*. Thus, Neisser (1978) could be quite correct in criticizing *typical* laboratory research on memory for exhibiting low ecological validity and neglecting critical aspects of memory. This does not mean that, *in principle*, laboratory research should exhibit those unfortunate features. Though it may require a bit of ingenuity in implementation, experimental manipulation and measurement of independent and dependent variables could still be exceptionally involving and full of impact for research. Laboratory research could pertain to issues participants hold "near and dear," and engage substantial degrees of motivation and affect, quite comparable to those found in many "everyday" circumstances.

Even more fundamentally, perhaps, the *generalized comparisons* between laboratory and real-life settings often seem unwarranted. That is because there exists no well-defined population of real-life circumstances that one could meaningfully characterize as a category, nor is there such a population of laboratory situations. Both "real life" and "laboratory" settings are immensely heterogeneous, with little to distinguish them from each other as separate classes of situations. The bored and apathetic "couch potato," the firefighter in a ferocious struggle to save lives, or the young bride on the eve of her honeymoon - all represent "real life" instances, yet the psychological situations they confront appear to

have very little in common. Thus, it seems rather risky to generalize about the accuracy rates of memory in field versus lab studies (Fisher et al. 1989). The question of where those rates are higher would seem extremely difficult to answer without knowing specifically which real-life and laboratory instances one has in mind. In short, the generalized dichotomy between real-life and laboratory research seems problematic on conceptual grounds. Hence, its implicit linkage to the storehouse—correspondence framework may hurt the new paradigm more than it can help.

A major substantive question concerns how the storehouse and the correspondence metaphors are interrelated. K&G contrast the two approaches as "competing" and note that they "reflect fundamentally different ways of thinking about memory." They also counsel against attempts to bridge the gap or merge the two, and prefer instead to "see their differences sharpened" (sect. 6.2, para. 4). Yet there is a sense in which the two metaphors are not only compatible but (also) mutually complementary, so that a full-fledged understanding of memory can hardly be achieved without considering them jointly.

The correspondence metaphor predominantly treats memory as a judgment about past events. The emphasis on accuracy neglects the process of *how* such a judgment was forged, focusing instead on its fit with some agreed-upon criterion. But the forging process should in fact matter a great deal, for not all judgments about past events merit the label of "memory." Specifically, judgments about the past can be made "on-line" rather than being memory-based. The former, strictly speaking, do not represent remembering in any meaningful sense.

To say this differently, though *all remembering entails some kind of judgment, not all judgments about past events constitute remembering*. In this regard, the classic "storehouse" metaphor seems to better capture the unique character of memory phenomena. The initial encoding episode, its storage and retrieval, and the issues of forgetting and interference — so central to the storehouse metaphor - are also central to the phenomena of memory *sui generis*. Even if the memory trace does not represent a specific and discrete recording of an event, even if it represents "diffuse and widespread modifications of the whole cognitive system, something in the system must change as a result of experience" (Craik 1983, p. 356). Only to the extent that this "something" plays a part in subsequent remembering performance can the phenomenon be properly described as one of memory.

There is the danger that isolated concentration on the correspondence metaphor may lose that critical nucleus of memory effects. One might rethink whether the crucial correspondence issue in memory research should be the correspondence between memory performance and an external criterion, or correspondence between memory performance at time 2 and the encoded event at time 1. The latter taps the fit between what we once encoded and now may or may not remember, whereas the former might refer to what we now constructed, but never actually encoded.

## Let's forget the everyday/laboratory controversy

Lia Kvavilashvili<sup>a</sup> and Judi Ellis<sup>b</sup>

<sup>a</sup>Division of Psychology, University of Hertfordshire, Hatfield, Herts AL10 9AB; <sup>b</sup>Department of Psychology, University of Reading, Reading RG6 6AL, United Kingdom, l.kvavilashvili@herts.ac.uk; xseliju@reading.ac.uk

**Abstract:** In contrast to its aims, Koriat & Goldsmith's article vividly demonstrates (1) the complementarity of ecological and traditional approaches and (2) the difficulty of characterising the growing diversity of memory research with a single set of distinctions. Moreover, the contrast between correspondence and storehouse metaphors is important enough

to stand alone without reference to an everyday/laboratory controversy, which is neither acute nor necessary.

The controversy about everyday versus laboratory-based research has a long history and is not confined to the study of memory (Kvavilashvili & Ellis 1995). However, as Koriat & Goldsmiths (K&G's) target article demonstrates, it is within memory research that this controversy has provoked extreme positions and elicited heated and recurrent discussions. It is therefore pertinent to ask what the state of this controversy was prior to the publication of this article and whether or not the article contributes to its resolution.

The latest round in the controversy undoubtedly began in 1978 (Neisser) and reached its peak in 1989 (Banaji & Crowder). It has since been in decline, probably because of the increased versatility of recent research practices, which make it difficult, if not impossible, to draw a clear distinction between the ecological and laboratory approaches to the study of memory. Indeed most researchers, irrespective of their orientation, admit that any tension between the two approaches is being gradually resolved in favour of peaceful coexistence and mutual benefit. This view has prevailed not only at several conferences (e.g., Kihlstrom 1994; Neisser 1988; Winograd 1991) but also among the participants of a debate in *American Psychologist* (January 1991) prompted by Banaji and Crowder's paper (see also Davies & Logie 1993). It would appear, however, that the primary aim of the current target article is to demonstrate that the controversy has not been resolved. By presenting new experimental data and, significantly, by introducing the correspondence metaphor (in opposition to the prevailing storehouse one) K&G will succeed in provoking a fresh debate. This seems inevitable, given their expressed, concluding desire to see the differences between the two approaches "sharpened and cultivated." Have the authors indeed managed to increase the breach between the two approaches to memory?

By emphasising the importance of variables such as memory property, report option, and test format, K&G have reformulated the debate on the everyday/laboratory controversy on more precise and rigorous grounds. Contrary to their final assertion, however, the target article reveals not only the complementarity of the two approaches but also the difficulty of finding a set of distinctions that can clearly differentiate them. For example, with respect to the recall-recognition paradox, K&G convincingly argue that the research setting (the "where aspect") may be the least important dimension in the everyday/laboratory controversy. What is important, however, is that variables identified in an everyday setting are exposed to empirical test. Thus, the superiority of recall over recognition, established in naturalistic studies, may be attributed to the use of accuracy measures rather than the research context per se. Clearly, however, accuracy and quantity measures complement each other and should not be viewed in isolation. For example, there are likely to be qualitative as well as quantitative differences between two persons who reveal 100% accuracy on a standard free-word recall test yet differ widely in the number of words they recall. Moreover, the superiority of recall over recognition has been amply demonstrated in laboratory studies using only quantitative measures (see, e.g., Tulving & Thomson 1973; Watkins & Tulving 1975). Clearly another important variable in both research contexts is the similarity between target items and foils in a recognition test (e.g., Bahrck & Bahrck 1964).

Finally, consider the "what," "where," and "how" dimensions of the everyday/laboratory controversy. K&G cogently point out that these dimensions do not provide a clearcut distinction between the two approaches. Instead, they offer a contrast between correspondence and storehouse metaphors which, they suggest, can account for some correlations between the three dimensions and clarify important aspects of the everyday/laboratory controversy. However, although their analysis of these metaphors is important with respect to the development of memory research, it does not clarify the fuzzy boundaries between the everyday and laboratory approaches. On the contrary, it explicitly reveals the difficulty in

drawing a clear line between these approaches. As K&G note, a great deal of correspondence-oriented research is conducted within the laboratory, much everyday research continues to use a storehouse metaphor, and there are clear signs of a shift "toward a correspondence-oriented metatheory . . . in a wide variety of contemporary approaches, including the reconstructive, attributional, ecological, functional, nonmediational, procedural, and connectionist approaches to memory" (sect. 3, para. 1).

In view of these important trends, it is reasonable to conclude that tension between the everyday and laboratory approaches in memory is neither acute nor necessary (cf. Kvavilashvili & Ellis 1995). Moreover, the contrast between a storehouse and a correspondence metaphor is, we suggest, sufficiently important to stand alone without reference to the everyday/laboratory controversy. This contrast is likely to prove more interesting theoretically than endless recurrent discussions about the merits and drawbacks of the everyday and laboratory approaches to memory. Thus it is our firm hope that in the near future we will be able to say "The everyday/laboratory controversy is dead, long live the correspondence/storehouse distinction!"

## Correspondence to the past: The essence of the archaeology metaphor

Steen F. Larsen

*Institute of Psychology, University of Aarhus, DK-8240 Risskov, Denmark.*  
steen@psy.aau.dk

**Abstract:** The correspondence view of memory is not a metaphor. However, correspondence is the essential feature of the archaeology metaphor, which harks back to Freud and Neisser. A modern version of this metaphor and some of its implications are briefly described. The archaeology metaphor integrates the idea of stored traces in a non-mechanistic framework.

The discussion of memory metaphors and metatheoretical assumptions initiated by Koriat & Goldsmith (K&G) is timely and potentially very useful, considering that the controversy over naturalistic versus laboratory approaches has partly stymied theoretical development in this area for more than a decade. K&G clearly point to a number of problems and limitations inherent in the storehouse metaphor, to which most traditional work on memory arguably conforms. Furthermore, they show convincingly that an alternative view, called the correspondence metaphor, leads to quite different research questions and data interpretations in cases like the "recall-recognition paradox" (sect. 5).

According to the correspondence view, memories should be evaluated by the degree to which they correspond to something that was the case in the past, that is, their accuracy or faithfulness as representations, rather than by the quantity of reproduced "items." Though not without problems of its own, I believe the focus on memories' correspondence to reality is an important change of priorities. However, as K&G seem to acknowledge in section 2.2, the correspondence view is not a metaphor at all. It does not confer any surplus meaning from its source domain to the understanding of memory, and it does not suggest any further properties of the memory system. Rather, it identifies a core feature to be included in an alternative metaphor.

As a useful alternative to the storehouse metaphor, can a full-fledged memory metaphor be found that embodies the correspondence criterion? I suggest that the *archaeology metaphor* may serve this purpose. I have already discussed this metaphor for memory in some detail (Larsen 1987). What follows is a brief summary of major points and a consideration of how the correspondence criterion and the real-life versus laboratory controversy appear through the prism of the archaeology metaphor.

The archaeology metaphor was used extensively by Freud (e.g., 1905/1953), who described himself as searching for pieces of the

past in his clients' mind from which to reconstruct their life experiences. If we eliminate the omniscient analyst (Freud) from the picture and let the rememberer do the reconstructive work, we come close to Neisser's (1967) comparison of memory to paleontology: "Out of a few stored bone chips, we remember a dinosaur" (p. 285). Nevertheless, the archaeology metaphor has been largely ignored in cognitive psychology (Roediger 1980). One reason may be that the metaphor was misunderstood by simply viewing archaeology as a hunt to uncover stored treasures. This "Indiana Jones version" of the metaphor is similar to the storehouse concept of memory, except for some romantic overtones. However, if a more scientific and modern view of archaeology is taken, the metaphor suggests a concept of memory that is remarkably close to K&G's correspondence view (cf. sect. 2.2) - without ignoring that traces from the past are indeed stored somehow.

To achieve correspondence between a present account and past reality is precisely the overarching aim of archaeology. As in memory, collecting items from the past only serves the purpose of constructing true descriptions to represent the past. As in memory, the number of items found is less important than their quality, that is, the accuracy with which they can point to significant aspects of the past. As in memory, the items themselves never come down to the present unchanged; they evince the wear-and-tear of intervening events - being corroded, scarred, fragmented, dislocated. As in memory, the context in which items are found is often decisive for their accurate interpretation. As in memory, the digging out of items disturbs this context irreversibly, and the change of context may even destroy the objects found. As in memory, there may be a mixture of concrete analogical and verbal remains, the latter possibly in a hard-to-understand language. And, as in memory, most archaeological objects were not laid down deliberately for posterity, though a number of monuments to the ruling classes (read: the central executive) do exist.

One shortcoming of the archaeology metaphor concerns cuing and search processes. Archaeological finds are almost never found in a pinch, when they are needed. However, professional excavations are carried out in a far more systematic manner than the romantic picture of the lucky find has it; the archeologist's knowledge about the way life was organized in the past suggests pretty well where to look. On the other hand, the continued importance of entirely serendipitous finds might call attention to the role of involuntary retrieval in memory.

The archaeology metaphor might seem to suggest that studies of memory must be carried out "in the field," that is, in real-life settings. Field excavation is indeed the primary method of archaeology. However, experimental work has gained a significant place in modern archaeology, for instance, studies of the traces remaining after the burning of reconstructed houses or the effects of particular patterns of use on tools and weapons. Thus, there is a place in the metaphor for laboratory studies. It might contain a useful lesson for the psychology of memory that such studies are always conducted to elucidate phenomena that have first been identified and described in the real-life context of the field.

Finally, in the archaeology metaphor the flexibility of the subject-controlled processes that K&G emphasize cannot be overlooked, because there is a human archeologist in charge who evaluates and narrates the meaning of the finds. This may have a ring of homunculus theory to it. But if we desire a system with a power to achieve correspondence to reality that equals the human mind, a mechanical metaphor will hardly do. K&G have not offered such a metaphor but have at least shown the need for it.

#### ACKNOWLEDGMENT

This commentary was written during a sabbatical stay at the Department of Psychology, University of Bristol, United Kingdom.

## Accuracy and quantity are poor measures of recall and recognition

Andrew R. Mayes, Rob van Eijk, and Patricia L Gooding

Department of Clinical Neurology, Royal Hallamshire Hospital, Sheffield, S10 2JF, United Kingdom, a.mayes@sheffield.ac.uk

**Abstract:** The value of accuracy and quantity as memory measures is assessed. It is argued that (1) accuracy does not measure correspondence (monitoring) because it ignores omissions and correct rejections, (2) quantity is confounded with monitoring in recall, and (3) in recognition, if targets and foils are unequal, both measures, even together, still ignore correct rejections.

We wish to focus on the kinds of memory measures that Koriat & Goldsmith (K&G) derive from storehouse metaphors and the correspondence "metaphor." They prefer these measures to others, and, in particular, those of signal detection theory (SDT). Accuracy is regarded as a measure that derives from notions of correspondence, whereas quantity derives from storehouse metaphors. We have two initial comments on this. First, accuracy ignores omissions. But if it is regarded as a measure of correspondence, then omissions are as important as false alarms. Memory is like perception, where accuracy is compromised as much by not seeing what is there as it is by seeing what is not there. Second, appropriate storehouse metaphors of episodic memory should suggest that adequacy of encoding, storage, and retrieval determines the completeness and accuracy of memory for experienced episodes, that is, how good the correspondence is. This suggests a quantitative measure of episodic memory that relates to correspondence and is a function of both kinds of error. SDT provides such a measure for recognition, but not for recall, as K&G argue. Our contention is that neither does their accuracy measure; nor at present, does any other recall measure.

K&G aim to analyse recall performance by assessing retention, monitoring, and control. For them, retention is target-item retrieval, monitoring is discrimination between target and foil items, and control is the giving or withholding of memory responses. Retention is measured by quantity (hit rate), monitoring with accuracy (the probability that a remembered item is a target) at different levels of confidence, and control by the level of confidence at which subjects make a free-recall response.

Hit rate and accuracy are poor measures of retrieval and monitoring in recall. Thus, hit rate in forced recall as a measure of retrieval is confounded with monitoring whenever monitoring fails. If a target is retrieved, subjects will miss it when they reject it and give a different response. When a foil is retrieved first, but the target is also retrievable, subjects will miss it if they wrongly accept the foil. The hit rate therefore is not a pure measure of retrieval. A problem with assessing monitoring in recall is that the experimenter does not know how many target items and how many "foil" items subjects generate when they try to recall. This is the reason that SDT cannot be used to provide a measure of monitoring in recall. The accuracy measure can still be computed, because it depends only on the produced hits and false alarms, but it is a poor measure of monitoring because it does not take the misses and the correct rejections into account. For example, if we know that subjects recall 12 targets and 3 "foils," then this does not tell us much about their monitoring performance. They may have generated 15 targets and 15 foils, giving them a monitoring hit rate of 0.80 and a monitoring false alarm rate of 0.20. This would give a  $d'$  value of 1.68. However, they may have generated 20 targets and 5 foils. In this case, their monitoring hit and false alarm rates would both be 0.60 and this would give a  $d'$  value of 0. The accuracy measure would be 0.80 in both these cases.

By our view, accuracy and quantity are equally inappropriate as measures of retention and monitoring in recognition. Furthermore, K&G use the term "retention" differently in recall and recognition. In recall it means retrieval, whereas in recognition it means memory strength. They also appear to be giving monitoring a different meaning when they argue that SDT cannot distinguish

between memory strength and monitoring. In this sense, monitoring effectiveness would be the ability to discriminate memory signals from noise. Performance in any recognition test presumably depends on subjects' retention and monitoring effectiveness in the above senses, and standard recognition tests (including the ones used by K&G) do not enable one to obtain separate measures for these two variables, whether one uses SDT or any other available memory measure (including accuracy). To do so would require being able to measure memory strength directly, uninfluenced by monitoring effectiveness. At present, neither K&G nor anyone else can do this.

K&G cannot do this any better than SDT because the quantity and accuracy measures are either equivalent to the sensitivity and response bias measures of SDT or they give misleading results, as we now show. If the numbers of targets and foils are equal on a Yes/No recognition test, then the hit rate ( $h$ ) and the false-alarm rate ( $fa$ ) can be computed from quantity ( $Q$ ) and the accuracy ( $A$ ) measures:

$$\left( h = Q, fa = \frac{1 - A}{A} \times Q \right)$$

Conversely, the quantity and the accuracy measures can be computed from the hit and the false alarms rates:

$$\left( Q = h, A = \frac{h}{h + fa} \right)$$

Therefore, in this case, the quantity and accuracy measures contain the same information as the hit and false alarm rates and hence, the sensitivity and the response bias measures of SDT. Which measures are appropriate depends on the situation. For example, accuracy may be important in a courtroom, but if you are interested in memory processes, then SDT is better, because sensitivity relates more directly to underlying memory processes than does accuracy or quantity, which are functions of both sensitivity and response bias.

If the number of targets and foils is unequal, then the false-alarm rate cannot be computed from the quantity and accuracy measures, hence they do not contain the same information as the sensitivity and the response bias. Quantity and accuracy together are not sensible measures in this case because they ignore correct rejections. This can be illustrated by making the number of foils that the recognition test contains very small. In this case, the number of false alarms must be small, and the accuracy measure will accordingly be close to 1. However, this does not mean that subjects' monitoring ability is good, because accuracy will remain close to 1 even if subjects are incapable of distinguishing between target and foils. We conclude that accuracy and quantity do not provide appropriate measures of either recall or recognition.

## The phenomenal object of memory and control processes

Giuliana Mazzoni

Department of Psychology, 50125 Florence, Italy, [gium@cesit1.unifi.it](mailto:gium@cesit1.unifi.it)

**Abstract:** This commentary deals with criteria for assigning truth values to memory contents. A parallel with perception shows how truth values can be assigned by considering subjects' beliefs about the truth state of the memory content. This topic is also relevant to the study of processes of control over retrieval.

Koriat & Goldsmith's (K&G's) correspondence view of memory entails concern with the truth value of memory content. The stress this places on accuracy forces us into deeper reflection about memory errors in research on everyday memory mechanisms and especially memory distortions (Loftus 1991).

We discuss whether the content of the real event is always the best information against which to judge the correctness of mem-

ory contents, or whether under some conditions subjective certainty about memory content might be a better candidate.

K&G mention the possibility of viewing errors as deviations from veridicality (sect. 2.2, para. 2) and discuss the problems connected with the attempt to obtain a holistic measure of accuracy (sect. 4.1). The same problems also affect analytic measures of accuracy, however, requiring us to understand better the criteria for deciding when an output is a error of commission and when it is a correct response. K&G acknowledge the existence of various degrees of deviation from reality and of different types of memory errors. However, memory content is true (or accurate) only when it faithfully represents the content of the real event. Whatever the degree of deviation, the comparison is necessarily with reality.

To explicate this point, K&G propose a parallel with perception, but the parallel also underscores a remarkable difference between memory and perception with respect to the meaning assigned to the correspondence between the content of reality and the content of the cognitive process.

In perception, the correspondence between what is perceived and what is out there in the real world is a way of understanding how perceptual mechanisms work and how they process raw visual information (e.g., Biederman 1987; Marr 1982). The object in the real world had also been called the "physical object," and the object of perception the "phenomenal object." The latter is often very different from the former, as shown clearly by visual illusions, but the phenomenal object has almost never been considered in itself perceptual error; the content of perception is usually the "true" perceptual datum.

In memory, in contrast, any deviation from the content of an event is considered a potential error, because it is the basis for assigning truth value. But should discrepancies from reality always be considered errors? Consider, for example, perceptually bistable figures (e.g., the Jastrow bistable figure). When presented for very short times, only one of the two possible interpretations wins, and subjects remember an incomplete figure that leaves out the parts that are inconsistent with the interpretation (Chambers & Reisberg 1992). Is this a memory error? Or is it rather a manifestation of the phenomenal object of memory?

Consider also the sizeable literature on phenomenal judgments of memory contents. One example is the distinction between "remember" and "know" judgments (Gardiner & Java 1993). Such judgments do not simply reflect confidence, but a mental reliving of the physical experience. Could these judgments not be considered marks of the "phenomenal object" of memory, thus representing "true memories"?

None of these phenomenological questions have answers here. It is important to note, however, the substantial impact that a deeper reflection on errors might have, not only on measures of accuracy, but also on theories about subjects' control over memory performance.

The role of subjects' monitoring and control over memory performance has been pointed out in metacognitive approaches to memory (Nelson & Narens 1990; Schneider & Pressley 1989), and K&G put together a convincing demonstration that memory accuracy is under strategic control. When the experimental procedure boosts accuracy, memory errors decrease to a very low level, and sometimes disappear (sect. 5.3.1), whereas in the quantity-oriented approach memory cannot be enhanced by similar procedures (e.g., Roediger & Payne 1985). These data show that the manipulation of commission errors is one of the crucial aspects of strategic regulation of memory performance. Confidence judgments are responsible for the decision to output the response and thus are responsible for accuracy, but confidence judgments are probably based on what is here called the "phenomenal object" of memory. Subjects are highly confident in what they believe to be true, and this does not necessarily correspond to what is true in reality. Hence, there are two types of correspondence that ought to be taken into account when studying control processes: the correspondence between the true content of reality and the memory representation; and the correspondence between the

memory representation and what the subject believes to be true. Correspondingly, there are also two ways to assign truth value to a memory content. It is interesting to note that the subjects' control operates according to the second type of correspondence, whereas the experimenter's control operates according to the first type of correspondence. This is a crucial discrepancy that must be considered in research on memory accuracy and on control processes involved in retrieval:

## False dichotomies and dead metaphors

Timothy P. McNamara

Department of Psychology, Vanderbilt University, Nashville, TN 37240.  
mcnamara@ctrvax.vanderbilt.edu

**Abstract:** Koriat & Goldsmith's thesis is provocative but has three problems: First, quantity and accuracy are not simply related, they are complementary. Second, the storehouse metaphor is not the driving force behind contemporary theories of memory and may not be viable. Third, the taxonomy is incomplete, leaving unclassified several extremely influential methods and measures, such as priming and response latency.

While reading Koriat & Goldsmith's (K&G's) target article, I was reminded of Cronbach's (1957) appeal to unite the Tight Little Island of experimental psychology and the Holy Roman Empire of correlational psychology, of Garner's (1972) analysis of applied and basic research, and of the ongoing controversy between information processing and connectionism. The history of our discipline is replete with "good guy-bad guy" controversies of this kind. These controversies and discussions of them play an essential role in the theoretical development of psychology. History tells us that they lead to new insights about the science and stimulate creative integration. K&G should be applauded for bringing the everyday-laboratory controversy to the bench and exploring its philosophical and methodological foundations. Their arguments, however, are not uniformly convincing.

K&G identify and successfully distinguish two alternative emphases in the measurement of memory, quantity versus accuracy. This insight is new, and it leads to a clear advancement in the methodology of memory research. The dissociation of "memory property" (quantity vs. accuracy), "report option" (free vs. forced choice), and "test format" (recall vs. recognition) in a common experimental design is an important contribution. It is worth noting, however, that K&G's results do not bode well for the everyday memory movement, inasmuch as findings thought to be unique to natural, social contexts of remembering can also be replicated in a sterile laboratory, using the "banal" task of memorizing a list of unrelated words.

However, in their zeal to distinguish between the storehouse and the correspondence metaphors, and to draw the causal arrow from these metaphors to mode of assessment, K&G have missed the deep complementarity between quantity and accuracy. This complementarity has roots in how quantity is assessed: Items are examined to determine which are recalled or recognized and which are not; at this level, the focus is solely on accuracy. At a more global level, accuracy is scaled by quantity: *Ceteris paribus*, who will the jury believe more, the witness who states that he remembers nothing (0% quantity; 100% accuracy), the witness who remembers the color of the assailant's shirt but nothing else (low quantity; high[?] accuracy), or the witness who constructs a detailed account of the event, including clothing, setting, the time of day, and so forth (high quantity; high[?] accuracy)? K&G acknowledge that quantity and accuracy are related (e.g., via tradeoffs) but don't seem to recognize how deep this relation really is.

K&G also fail to convince us that the storehouse metaphor is anything but dead. Conspicuous in their absence are references to contemporary theories of memory, such as ACT\* (Anderson 1983), SAM (Gillund & Shiffrin 1984), and TODAM (Murdock 1982), to

name just a few. These theories are founded on and evaluated against data produced by quantity-oriented research, such as learning lists of words or sentences. However, they are concerned with mental representations and processes, and their reach extends to domains as complex as language comprehension, problem solving, and human decision making. Indeed, the articles that K&G cite as exemplars of the storehouse metaphor represent an approach to memory that is striking in its emphasis on functional relations and its lack of attention to cognitive mechanisms. This approach might have been influential in the past, but it is not paradigmatic and may not even be viable.

K&G's taxonomy also left me scratching my head about how to classify associative or repetition priming, or, for that matter, any measure of processing time. It is hard to see how these methods and the uses to which they have been put are motivated by either a storehouse or a correspondence metaphor, or how they sort as quantity or as accuracy measures. When one finds, for example, that "lion" primes "tiger" in lexical decision (which is a form of memory retrieval), what has one learned? Certainly nothing about how big memory is. One has learned that a relation (i.e., association) exists in memory. This relation, however, has no embodiment in the world; to the extent "correspondence" exists, it does so only very abstractly.

My reluctance to embrace Koriat & Goldsmith's thesis can be traced, in part, to an absence of conflict: when I survey my own research, I see that most of it has been of the correspondence type, very little has been of the traditional storehouse type, yet all of it has been informed and enlightened by empirical and theoretical developments in quantity-oriented research. Perhaps I am the victim of repressed anxiety, blind to the intellectual battles raging in my subconscious, but I doubt it. I am moved instead by a different metaphor: when it comes to pulling chariots, two horses are better than one.

## Remembering as doing

Ulric Neisser

Department of Psychology, Emory University, Atlanta GA 30322.  
neisser@fsl.psy.emory.edu

**Abstract:** Koriat & Goldsmith are right in their claim that the "ecological" and "traditional" approaches to memory rely on different metaphors. But the underlying ecological metaphor is not *correspondence* (which in any case is not a metaphorical notion): it is *action*. Remembering is a kind of doing; like most other forms of action it is purposive, personal, and particular.

Koriat & Goldsmith (K&G) have made a major contribution. Their insight makes it possible to see the dispute between "ecological" and "traditional" approaches to memory (Neisser & Winograd 1988) in a new and clearer light. Although differences of method are often involved, the most fundamental difference - e.g., between my views and those of Banaji and Crowder (1989) - does not concern how research should be controlled or where it should be conducted; it concerns how we think about memory itself. There are indeed two contrasting "memory metaphors." But in my view the target article gets only one of those metaphors right; the other requires further analysis.

To assert with Locke that memory "is as it were the storehouse of our ideas" (sect. 2.1) is indeed to use a metaphor. Just as we may put things into a storehouse on one occasion and take them out again later, so too we put our experiences "into memory" at time 1 and (one hopes) retrieve them again at time 2. This classical *storehouse metaphor* is actually just one version of the fundamental *containment* schema described by Mark Johnson (1987; see also Lakoff 1987; Mandler 1992). It has been very productive, generating a host of interesting questions for memory research. Like other metaphors based on the containment schema, it is intuitively plausible: everyone understands putting objects into

containers and taking them out again. Despite its surface plausibility, however, everyone is also aware that it is *only* a metaphor. Life experiences are not really objects; memory is not really a storehouse, nor is it any other kind of physical container.

K&G contrast the storehouse metaphor with another way of thinking about memory, in which "the basic criterion for evaluating memory is not the quantity of items remaining in store, but rather, the *correspondence* between what the person reports and what actually happened" (sect. 2.2, para. 1). They call this the *correspondence metaphor*, and regard it as the basic intuition that underlies ecological approaches to the study of memory. But while there is no doubt that issues of accuracy and correspondence are important in these contexts, this way of describing the situation is misleading. Unlike "storehouse," correspondence is not a metaphor. It is better thought of as a goal at which remembering may (sometimes) aim, or as a criterion by which the adequacy of (some) recollections may be judged. When the witness says the perpetrator had a gun, when the patient says she was raped by her father, when the experimental subject says that he first learned about the *Challenger* disaster from TV, the most important question is the one about correspondence: Are they right? That question is in no way metaphorical; it concerns matters of fact.

To claim that something is a matter of fact is not to claim that it is easily assessed. K&G are quite right to emphasize the difficulty of measuring such "correspondences," that is, between what really happened at time 1 and how someone remembers it at time 2. Even the comparatively simple task of assessing the correspondence between two verbal accounts of the same event by the same person can pose substantial problems. The scoring system that we devised to compare two accounts of hearing the news about *Challenger* (Neisser & Harsch 1992) had to be modified even for the very similar task of comparing two accounts of one's experiences during an earthquake (Neisser et al., in press).<sup>1</sup> In some cases, the assessment of correspondence and accuracy can become so difficult that no resolution is possible at all (Edwards & Potter 1992). But this does not mean that "correspondence" is a metaphorical notion; it only means that some factual questions must remain unanswered.

There is another, perhaps more obvious, problem with the claim that the ecological study of memory revolves around "correspondence." Not all memory-based behaviors focus on accurate reproduction; other dimensions of the performance may be much more important. The aim in telling a joke, for example, is to tell it *effectively*; whether you tell it just as you heard it is of no consequence at all. A singer of epic tales "repeats" a familiar story, but not with a view to reproducing some prior performance; rather, his intent is to impress and entertain his present audience (Lord 1960; Rubin 1995). Actors do indeed memorize their lines, but getting the words right is the least, most insignificant part of their task. Experimental subjects tend to focus on accuracy when they recall a story for the benefit of the experimenter, but other dimensions become more important when they discuss the same story with a peer (Hyman 1994).

Those of us who are interested in these forms of memory do indeed have metaphorical ways of thinking about them, but the underlying referent is not correspondence. Rather, it is *action*. Remembering is a kind of doing. Like other kinds of doing, it is purposive, personal, and particular:

- (1) It is *purposive* because it is done with a specific goal in mind; often that goal is to tell the truth about some past event, but on other occasions it may be to entertain, to impress, or to reassure.
- (2) It is *personal* because it is done by a specific individual and bears the stamp of that individual's characteristic way of doing and telling.
- (3) It is *particular* because it is done on a specific occasion, in a way that reflects the particular opportunities and demands that the occasion may afford.

Different theorists prefer different versions of the action metaphor. Following Bartlett (1932), I myself have often metaphorically described memories as *constructions*, that is, as prod-

ucts that are skillfully built from available parts to serve specific purposes (e.g., Neisser 1967). Because they are constructions rather than copies, they can often be seriously mistaken even when the individual is explicitly aiming at correspondence. This metaphor still makes sense to me, even though I have long since given up constructivism in other areas. (Visual perception, for example, is better thought of as "direct" than as "constructive"; Neisser 1976). But still it is only a metaphor: There is no point to asking whether memory *is* or *is not* constructive (Brewer 1992), as if an empirical question were at stake.

Many other theorists have also described memory in terms of purposive action, though the specific forms of the metaphor vary widely. Michael Ross (1989) speaks of rememberers as drawing inferences from implicit theories; John Anderson (1991) deduces characteristics of memory from general principles of adaptive behavior; narrativists like David Gergen (1994) regard all remembering as nothing but storytelling. What these theorists do *not* do — the metaphor they do *not* use — is to regard memory as a place for passive storage. Thus, Koriat & Goldsmith are quite right to assert that the traditional and ecological approaches to memory are based on different metaphors. On one side is the notion that memory is a storehouse; on the other, the notion that remembering is a form of purposeful action.

#### NOTE

1. As the target article points out (sect. 5.1), memory is especially difficult to assess when subjects are free to choose the level of generality at which they describe some past event. This became clear in the Emory University study of memory for a semester-long seminar, which K&G cite as "Neisser (1988b)." Although I did describe the Emory study briefly in that paper, it was in fact a collaboration with Ira Hyman, Nicole Harsch, and JoNell Usher (Neisser et al., in preparation). Other data from the same study are reported in Hyman and Neisser (1992).

## Metacognition, metaphors, and the measurement of human memory

Thomas O. Nelson

Department of Psychology, University of Maryland, College Park, MD 20742.  
tnelson@glue.umd.edu

**Abstract:** Investigations of metacognition — and also the application of the storehouse and correspondence metaphors — seem as appropriate for laboratory research as for naturalistic research. In terms of measurement, the only quantitative difference between the "input-bound percent correct" and "output-bound percent correct" is the inclusion versus exclusion (respectively) of omission errors in the denominator of the percentages.

I comment here on three of the topics that the target article by Koriat and Goldsmith (K&G) focused on: metacognition, metaphors, and the measurement of human memory.

**Metacognition.** K&G highlight the impact metacognitive processing has on memory performance and make many useful remarks about both the practical and theoretical importance of metacognition (e.g., in sects. 4.2.2.2, 5.3.1, and 5.3.2). The findings about retrieval described by K&G fit exceptionally well into a theoretical framework for metamemory (see Nelson & Narens 1990, particularly their Fig. 5). However, I believe that subject-controlled (active) metacognition occurs not only in naturalistic research *but also* in laboratory research on memory (even when the effects of metacognition are not observed because the investigator tries to suppress the organism's active processing), as emphasized by Nelson and Narens (1994). Metacognition is much too relevant to restrict it to the domain of naturalistic investigations; systematic investigations in the laboratory also can give us useful information about how to characterize the role of metacognition in memory performance (for many examples, see Metcalfe & Shimamura 1994; Nelson 1992; 1996).

**Metaphors.** The correspondence metaphor and the storehouse metaphor emphasize different aspects of memory, are not mutu-

ally exclusive, and either (or both) of them can be applied to laboratory investigations as well as to naturalistic investigations. I believe that investigators should use whichever metaphors work best (following the long-standing use of metaphors for memory; see Carruthers 1990, Ch. 1), that different investigators will legitimately use different metaphors, and that in contrast to the substantial achievements in philosophy of science for evaluating theories, no satisfactory method is available for evaluating metaphors and analogies (Lachman 1960, p. 127). Moreover, metaphors are necessarily both right *and* wrong (for elaboration, see the distinctions between positive, negative, and neutral analogies in Hesse 1966). I accordingly believe that researchers should devote effort not to "further development of the metaphor" (final paragraph of sect. 6.1) but rather to the development of theory.

I cannot agree that there are "memory measures that follow uniquely from the correspondence metaphor" (sect. 3, para. 11). Rather, a given memory measure could follow from many kinds of metaphors, and many memory measures (as defined by the particular reduction sentences stipulated by investigators when operationalizing the terms in their theories) could follow from a given metaphor.

**Measurement.** The key distinction between what K&G refer to as a "quantity measure" and "accuracy measure" in terms of the input-bound percentage correct and the output-bound percentage correct (sect. 4.2, para. 2 and sect. 5.2, para. 2) can be crystallized by noticing that the difference is solely in terms of the inclusion versus exclusion (respectively) of omission errors in the denominator of those percentages, as shown here in Figure 1. Although K&G's interpretation of that distinction is interesting, we should not overlook exactly what the distinction is. Moreover, the quantification of memory performance - regardless of whether it occurs in terms of the number correct, amount of distortion, latency of response, or whatever - is (and should remain) a general goal for both the laboratory and naturalistic research.

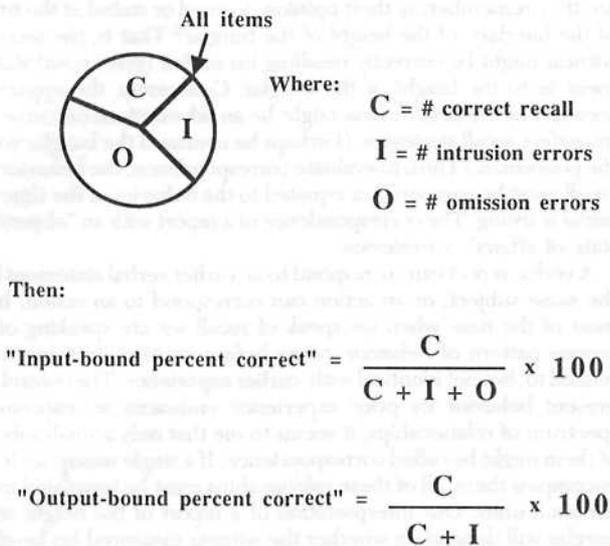


Figure 1 (Nelson). The only difference between what Koriat and Goldsmith (K&G) refer to as the input-bound percent correct (which K&G also refer to as "quantity" in para. 2 of sect. 5.2 and which analogously in the artificial-intelligence literature on case-based reasoning systems is referred to as "recall" in terms of the percentage of relevant cases that are retrieved) and the output-bound percent correct (which K&G refer to also as "accuracy" in para. 2 of sect. 5.2 and which analogously in the artificial-intelligence literature is referred to as "precision" in terms of the percentage of cases retrieved that are relevant) is shown here to be the presence versus absence of omission errors in the denominator of the percentages.

**Conclusions.** Although I applaud much of K&G's target article, I disagree with their final paragraph. Instead of cultivating different treatments of memory across the domains of laboratory and naturalistic research, we should generally try to close the gap between those two domains. In terms of the analogy in the final sentence of K&G's article ("each horse will be able to draw the chariot of science as far and as last as it can"), progress seems to be less likely if two horses pull in different directions; directions of pulling that are too different may even be counterproductive and could pull apart the metaphorical chariot. Rather, progress seems more likely when the two horses' efforts are additive. Moreover, if cross-fertilization occurs across the naturalistic and laboratory domains, the overall outcome across the two domains could be synergistic.

I wish to close this commentary by highlighting Koriat & Goldsmith's conclusion that "[a]n important challenge, then, is to develop ways of making the contribution of metamemory processes explicit in the evaluation of memory performance" (sect. 5.3.2, para. 4). This is a challenge that is likely to keep all of us busy for quite some time, both in the laboratory domain and in the naturalistic domain.

ACKNOWLEDGMENTS

Preparation of this article was supported by Grant R01-MH32205 and a Career Development Award (K05-MH1075) from the National Institute of Mental Health.

Beyond the correspondence metaphor:  
When accuracy cannot be assessed

Ian R. Newby and Michael Ross

Department of Psychology, University of Waterloo. Waterloo, Ontario, Canada N2L 3G1. irnewby@watarts.uwaterloo.ca; mross@watarts.uwaterloo.ca

**Abstract:** Koriat & Goldsmith propose that the correspondence metaphor captures the essence of everyday memory research. We suggest that correspondence is often not at issue because objective assessments of everyday events are frequently lacking. In these cases, other questions arise, such as how individuals evaluate the validity of memories and the significance they attach to those evaluations.

Koriat & Goldsmith's (K&G's) analysis of the differences and similarities between laboratory and everyday memory research offers the promise of theoretical and empirical advancement, as well as a rapprochement between the two traditions. We agree with much of what they say. In this commentary, we focus on the challenge that the study of everyday memory poses for the correspondence metaphor: correspondence is often difficult to assess and accuracy is not the sole or most important aspect of many everyday memories.

The correspondence metaphor assumes that researchers can access the target of a memory and "objectively" evaluate the accuracy of recall. In contrast, many everyday memories are difficult to verify. People may experience an event differently, an issue K&G consider briefly in their discussion of the *Rashomon* effect (sect. 4.1, para. 11). How people perceive a complex stimulus depends (in part) on their prior knowledge, expectations, and so forth. Individuals may experience the same event quite differently (Hastorf & Cantril 1954). Consequently, it is not clear what the test of correspondence should be. Perhaps researchers should evaluate memory against an individual's initial representation of the event, rather than against the supposed objective stimulus. After all, we cannot ask more of memory than that recollections reflect the person's original reality; otherwise, we confuse differences in memory with differences in perception. [See Dennett & Kinsbourne: "Time and the observer" BBS 15(2) 1992.] It is often difficult, however, for investigators to measure accuracy by the degree of correspondence between the initial

representation and a recollection. Researchers of everyday memory typically examine people's recollections of their past experiences and lack access to people's original representation of an episode (Ross, in press).

Moreover, there is a plethora of events for which objective accounts do not exist, and yet the memories are of psychological import. For example, Ross and Holmberg (1990; 1992) studied the vividness of married couples' memories for their first date, a shared vacation, and a recent argument. It was obviously impossible for the researchers to verify the accuracy of these memories. Nonetheless, the data are informative. Wives reported significantly clearer memories of the target events than did their husbands. Both spouses associated greater vividness with greater accuracy and assumed the superiority of women's memories. The researchers speculated that differences in memory quality might have important consequences for a relationship: "If a husband believes that he can recall the 1969 World Series more vividly than he can remember significant events in his own marriage, his wife, and indeed he himself, may well wonder if their relationship is not all that important to him." (p. 602) Thus, beyond the question of accuracy, qualities of memories such as vividness can play a significant role in everyday memory and therefore in everyday life.

Next, consider an example of rather bizarre memories. Newman and Baumeister (in press) investigated UFO (Unidentified Flying Object) abduction accounts. They found that a surprising number of people report having been abducted by extraterrestrial beings. Newman and Baumeister questioned the accuracy of these accounts and offered a psychological explanation for how false memories of UFO abductions could be created and maintained. Given that an abduction account is based on *some* event - Newman and Baumeister seem to accept this - it is an event for which an objective account does not exist. Thus, judgments of memorial accuracy must be based on indirect evidence, such as qualities of the recollections and characteristics of the rememberers.

Ross (in press) described criteria that individuals use when judging the validity of their own or other people's memories. These include: (1) source characteristics (who the rememberer is), (2) context of the recall, (3) memory qualities (e.g., vividness), (4) presumed memorability of the event, (5) internal consistency of the memory, (6) reliability of the memory, (7) congruence of the recollection with other knowledge and experiences, and (8) consensus - whether other people remember the episode in the same way. Ross and Newby (in press) use these criteria to explain why the abductees believe that their memories are accurate, whereas Newman and Baumeister do not. The beliefs of skeptics and abductees could both be supported if the parties used different truth criteria and/or weighted the criteria differently. For example, abductees tend to provide vivid accounts of their alleged abduction. The abductees may well take their ability to report details as evidence of the veracity of their memories. Newman and Baumeister acknowledged this vividness but did not admit it as evidence for the validity of the accounts. They emphasized other criteria on which the accounts fare less well.

Often, the accuracy of everyday memories cannot be assessed in a direct and objective way. We have described two types of events, one common and one exotic, in which this is the case. By focusing on the evaluation of accuracy, Koriat & Goldsmith risk excluding important aspects of everyday memory, such as the criteria people use to assess accuracy and the significance that they attach to their evaluations. If researchers limited their study to objectively verifiable memories, they would ignore many fascinating psychological phenomena and questions of theoretical consequence.

## Operationalizing "correspondence"

David C. Palmer

Department of Psychology, Smith College, Northampton, MA 01063.  
dcpalmer@science.smith.edu

**Abstract:** The research guided by the correspondence metaphor is lauded for its emphasis on functional analysis, but the term "correspondence" itself needs clarification. Of the two terms in the relationship, only one is well defined. It is suggested that behavior at acquisition needs to be analyzed and that molecular principles from the learning laboratory might be useful in doing so.

As a behaviorist, I was encouraged by the extent to which the target article attempts an experimental analysis of behavior with little, if any, dependence on hypothetical constructs. The research cited by Koriat & Goldsmith (K&G) moves us away from a preoccupation with inferred structures to a concern for the variables that control the behavior of remembering. It is true that many of these variables are covert activities of the subject - self-prompting, probing, rehearsing, editing, and evaluating - and are thus out of reach of tight experimental control, but at least they are behaviors, and presumably subject to the same independent variables as overt behavior. Consequently, our interpretations of these activities can be guided by principles of behavior derived from well-controlled laboratory investigations.

However, lest the praise of a behaviorist be the kiss of death to a new proposal in cognitive science, let me caution that I was also struck by the magnitude of the task facing the researcher working under the aegis of the new metaphor in its current form. The central term of the metaphor, correspondence, needs clarification. What does it mean to say that a verbal report "corresponds" to a past state of affairs? If one witness reports that a burglar was 6' tall, another that he was 5' 10", there is an obvious sense in which the first report might be more accurate than the second; nevertheless, as mnemonic behavior, the second report might be more accurate than the first. Are they remembering the height of the burglar, or are they remembering their opinion, formed or stated at the time of the burglary, of the height of the burglar? That is, the second witness might be correctly recalling his earlier (erroneous) statement as to the height of the burglar. Conversely, the apparent accuracy of the first witness might be an adventitious outcome of imperfect recall strategies. (Perhaps he confused the burglar with the policeman.) Thus, to evaluate correspondence, the behavior at recall must be compared or equated to the behavior at the time of initial learning. The correspondence of a report with an "objective state of affairs" is irrelevant.

A verbal report can correspond to an earlier verbal statement by the same subject, or an action can correspond to an action, but most of the time when we speak of recall we are speaking of a unique pattern of behavior, never before emitted, in some way related to, but not identical with, earlier experience. The control of present behavior by prior experience embraces an enormous spectrum of relationships; it seems to me that only a small subset of them might be called correspondence. If a single metaphor is to encompass them, all of these relationships must be translated into common units. Our interpretation of a report of the height of a burglar will depend on whether the witness measured his height, was told his height, asserted an opinion as to his height upon seeing him, saw him pass through a doorway of known height, merely glanced at him, saw him surrounded by unusually short people, saw him in a trapezoidal room, and so on. As jurors, we would care only about correspondence with an external state of affairs, but as students of memory, we care about the relationship between past experience and current behavior.

In my view, to accomplish this work of translation, the correspondence metaphor needs to be supplemented by a consideration of principles of learning. No theory of memory can be divorced from a theory of learning: The probability that a response will correspond to an earlier instance of behavior depends in large part on the strength of the behavior to begin with. Moreover, these

principles might provide the analytical tools necessary to bridge the gap between those phenomena that we can control experimentally and those we most wish to understand. The problems facing the student of memory are well known: Human behavior is a complex function of personal history and genetic endowment, neither of which is available in any detail to the researcher; ethical constraints on research with humans limit our control of historical and motivational variables; and, as noted above, many activities of fundamental importance are covert or below the threshold of observation with current technology. In short, achieving complete control of our subject matter is out of the question. Because principles of learning are relatively molecular, they provide a tool kit for guiding our inferences about memory phenomena where adequate controls are impossible (e.g., Donahoe & Palmer 1994; Palmer 1991). Although it requires reaching across paradigms, I can think of no reason why the principles of learning should not be exploited by researchers guided by the correspondence metaphor to evaluate the relationship between behavior at acquisition and recall.

### Amnesia and metamemory demonstrate the importance of both metaphors

Bennett L. Schwartz

*Department of Psychology, Florida International University, Miami, FL 33199. schwartzb@servax.fiu.edu*

**Abstract:** The correspondence metaphor is useful in developing functional models of memory. However, the storehouse metaphor is still useful in developing structural and process models of memory. Traditional research techniques explore the structure of memory; everyday techniques explore the function of memory. We illustrate this point with two examples: amnesia and metamemory. In each phenomenon, both metaphors are useful.

Koriat & Goldsmith (K&G) are right that the correspondence metaphor is a useful tool in generating functional models of memory. However, the storehouse model still has currency, particularly in the development of structural or process-oriented models of memory. Two examples to demonstrate this point: the differences between frontal and hippocampal amnesia and dissociations between memory and metamemory. Moreover, as K&G suggest, learning and remembering involve a complex interaction of memory and metamemory processes. As of now, the only method to study memory in and of itself is to hold metamemory constant. Traditional laboratory techniques can accordingly be defended.

Consider the following situation. A neuropsychologist is confronted with two amnesic patients. The first is painfully aware of her memory loss and does everything she can do to compensate for this loss. She seeks therapies, tries various mnemonic strategies, and writes things down, but it is all in vain. There is little she can do to regain her former memorability. The second patient is blissfully unaware of his memory loss. He confabulates, confuses one person with another, and has lost all track of time. Although these two patients may exhibit a similar percentage correct in a cued-recall test, it is clear that the form of amnesia is markedly different. Neuropsychological data suggest that the first amnesic is suffering from temporal or hippocampal damage, whereas the second is suffering from frontal damage (Moscovitch 1994).

The frontal amnesic may have no problem producing responses when memory cues are provided, but all of these responses may be commission errors. There is therefore little correspondence between the output of the memory system and the experienced events. In discussions of frontal amnesics, K&G's correspondence metaphor may be appropriate. The hippocampal patient, however, outputs almost nothing. Hers is a "storehouse" problem. Events and experiences in her life are not accessible to retrieval (or were never stored in the first place). It is not in the correspondence

between retrieval and reality that the hippocampal amnesic fails, it is in encoding/retrieving the memory trace. For the hippocampal amnesic, then, it is the storehouse metaphor that better captures the deficit.

K&G argue that the storehouse metaphor fails because it does not encourage researchers to examine important components of learning and remembering such as monitoring and control (metamemory). In contrast, the correspondence metaphor encourages research on metamemory because of its importance for memory accuracy. As a metamemory researcher, I endorse the importance of accuracy-based studies of memory. When memory performance is considered as a product of an interacting memory/metamemory system, the correspondence metaphor spawns useful functional models, appropriate for research in the "everyday" tradition. However, in developing structural or process models, it is crucial to distinguish between memory and metamemory components because they may influence memory performance in different ways. It is therefore important to study one while holding the other constant. When one holds metamemory constant, structural models based on the storehouse metaphor are appropriate in investigating "pure" memory (or "ecphory"; see Tulving 1983). When memory variables are held constant, one can look at the independent effects of monitoring and control functions (or "conversion," in Tulving's terminology).

Research has found functional dissociations between memory and metamemory in both amnesics and normals. Shimamura and Squire (1986) looked at feeling-of-knowing judgments in amnesic patients. Feeling of knowing is defined as a judgment of future recognition for an unrecalled item. Shimamura and Squire found that hippocampal amnesics showed very poor memory performance, but the accuracy of their judgments was above chance and did not differ from that of normal controls. Korsakoff's amnesics, who show both frontal and hippocampal damage, showed poor memory performance, and their accuracy did not differ from chance. Thus one finds a neuropsychologically based dissociation between memory and metamemory. Several studies have documented dissociations between memory and metamemory in normals (see Schwartz, 1994, for a review). For example, Koriat (1995) looked at two different sets of general-information questions: a difficult set and a deceptive set. Recall performance was equivalent, but accuracy in a feeling-of-knowing task was above chance for the difficult set and below chance for the deceptive set. Based on such findings, most metamemory researchers consider metamemory and memory to be separate and dissociable processes.

Now consider a "traditional" laboratory experiment. Participants are presented with a list of paired associates. After study and a retention interval, a four-alternative, forced-choice recognition test is given. K&G argue that this gives us only an input-bound measure, and that the ability to monitor and control has been reduced to almost zero. K&G rightly criticize these kinds of laboratory experiments because for years they were done without ever considering the roles of monitoring and control. Recent attention to metamemory has begun to alter that situation. However, these traditional experiments do allow us to consider memory unconfounded by another cognitive process: metamemory. Only when metamemory processes are held constant can we examine what variables affect basic memory processes. And correspondingly, we must hold memory processes constant to examine metamemory.

#### ACKNOWLEDGMENT

I would like to thank Ron Fisher and Robert Bjork for useful discussion concerning the ideas in this commentary.

## Classical antecedents for modern metaphors for memory<sup>1</sup>

Jocelyn Penny Small

Alexander Library, Rutgers University, New Brunswick, NJ 08903.  
jpsmall@gandalf.rutgers.edu

**Abstract:** Classical antiquity provides not just the storehouse metaphor, which postdates Plato, but also parts of the correspondence metaphor. In the fifth century B.C., Thucydides (1.22) considered the role of gist and accuracy in writing history, and Aristotle (*Poetics* 1451b, 1460b 8-11) offered an explanation. Finally, the Greek for truth (*aletheia*) means "that which is not forgotten."

Classical antiquity addressed certain issues in ways that still affect us. The metaphor of the storehouse for memory goes back to that time, though not as early as Plato. He compared memory to writing on wax (*Theaetetus* 191ff.) and reserved the image of the storehouse for writing alone (*Phaedrus* 276d1; cf. Xenophon, *Memorabilia* 1.6.14). The image of the aviary (*Theaetetus* 197c) is not quite the same thing, because, strictly speaking, the Greek and Latin comparisons are not to something as mundane as a "storehouse" but to a "treasury" (*thesauros* and *treasaurus* respectively), with a concomitant emphasis less on quantity and more on selectivity. The Greeks and Romans did not think of using written words to find other written words. Instead, they used one of the tools they already had: memory, which became *the* classical means of cognitively organizing and, most significantly, retrieving, words. Hence memory, the means of finding the words already stored in treasuries, itself came to be "the treasury of all things" (Cicero, *De Oratore* 1.18).

While no classical source uses the term "correspondence," classical antecedents do exist for some of its components: "gist" and "accuracy." The earliest discussion is not by a philosopher, but an historian, an ancient equivalent of the ecologist. Thucydides (1.22), writing in the fifth century B.C. about the Peloponnesian War between Athens and Sparta, describes his methodology:

As to the speeches that were made by different men . . . it has been difficult to recall with strict *accuracy* the words actually spoken. . . . Therefore the speeches are given in the language in which, as it seemed to me, the several speakers would express, on the subjects under consideration, the sentiments most befitting the occasion, though at the same time I have adhered as closely as possible to the *general sense* [the gist] of what was actually said. But as to the facts of the occurrences of the war, I have thought it my duty to give them, not as ascertained from any chance informant nor as seemed to me probable, but only after investigating with the greatest possible accuracy each detail, (emphasis added)

When Thucydides says that in certain cases like the speeches accuracy is not possible, we tend to agree, since he lived before the invention of shorthand (around the time of Cicero) and such modern amenities as the tape recorder. When Thucydides makes the further distinction that giving not the gist, as we might say, but what the "speakers would express" is appropriate, we are squarely in the midst of one of the current controversies over the nature of memory.<sup>2</sup>

Furthermore, the gist, according to Aristotle, can be more telling than accuracy. While we consider mimetic fidelity better - either it happened this way or it did not - Aristotle in the *Poetics* (1460b 8-11) reverses our judgment by putting "as they ought to be" at the top of his three levels of imitation, and "as they were or are" at the bottom. In between is "as they are said to be or seem." Aristotle also says (*Poetics* 1451b) that "The real difference [between history and poetry] is this, that one tells what happened and the other what might happen. For this reason poetry is something more philosophical and serious than history, because poetry tends to give general truths while history gives particular facts." Barclay and Wellman (1986, p. 101) echo Aristotle: "What one remembers then is, at least in part, what could have happened or should have happened in one's life." By classical standards, then, Neisser's (1981, p. 159) conclusion about John Dean's testimony makes

sense: "Dean's consistency was deeper; he recalled the theme of a whole series of conversations and expressed it in different events. . . . Episodes are not the only kinds of facts."

In another article with a Thucydidean title, Barclay (1988, p. 291) uses the same terms as Thucydides in the same way: Our claim is that overconfidence in the accuracy of autobiographical memories is due to the fact that many "events" could have happened as remembered because such events are consistent with the theme or gist of one's life - they are truthful but inaccurate recollections. This gist tends to remain fairly stable over time because it is derived from a relatively stable sense of self.

Barclay brings up two "classical" issues. First, as Thucydides and Aristotle implied: the opposition is not simply between the quantity of the memories (storehouse) and their accuracy (correspondence) but within the correspondence metaphor, between what kinds of memories are subject to being remembered with specific details and what kinds are remembered only in a general sense. Second, Barclay raises the issue of "truth." Here classical Greek is curious, for its word for truth, "*aletheia*," literally means "that which is not forgotten," as in the river Lethe (forgetfulness) in the Underworld. Hence whatever is remembered is "true": a sobering thought for all memory researchers.<sup>3</sup>

### NOTES

1. The translations are adapted from the *Loeb Classical Library* editions published by Harvard University Press.
2. On Thucydides' reliability, one of the key issues facing memory researchers, consider what Green (1993) says: "If we were confronted with a history of the Second World War written by an ex-general dismissed for incompetence, who in his account did a hatchet job on the man responsible for his dismissal, omitted a number of key events, refused to cite sources or authorities (so that his reiterated claims to objectivity had to be taken on trust, except when other evidence became available, which often contradicted him) and put words into the mouths of his characters as he felt appropriate — what would we think about such a work and such a man, however well he wrote?"
3. It may be significant that Socrates in Plato (*Cratylus* 421b) derives *aletheia* from "*theia ale*," or "divine wandering." Neither the Greeks nor the Romans were especially good at etymologies.

## Direct remembering and the correspondence metaphor

K. Geoffrey White

Department of Psychology, University of Otago, Dunedin, New Zealand.  
kgwhite@otago.ac.nz

**Abstract:** The correspondence view is consistent with a theory of direct remembering that assumes continuity between perception and memory. Two implications of direct remembering for correspondence are suggested. It is assumed that forgetting is exponential, and that remembering at one time is independent of factors influencing remembering at another. Elaboration of the correspondence view in the same terms as perception offers a novel approach to the study of memory.

The correspondence view proposed by Koriat & Goldsmith (K&G) sees memory as a reflection of the environment. This emphasis is generally consistent with a theory of direct remembering (TDR), which follows Gibson's (1979) claim that perception is direct (White 1991). In TDR, memory and perception are continuous and the same discrimination principles apply to both. Environmental information at the time of retrieval allows direct perception of the remembered event. The event is not stored but perceived directly, albeit at a temporal distance. Through active attempts to seek information in prior learning, the system is tuned to resonate to information available at retrieval. The information includes the retrieval context and the signals that context provides to initiate remembering, as well as the temporally distant event. Just as discriminations of spatially distant events are difficult, so too are discriminations of temporally distant events. Temporal

distance is thus a property of the event, just as other attributes are. Temporal distance can render an event difficult to remember; as a result, it can also facilitate the discrimination of recency.

Two corollaries of TDR have interesting implications for the correspondence view. Both follow from the fundamental assumption that perception and memory are continuous (see sects. 2.2 & 3).

The first implication is that the reduction in veridicality or accuracy over a certain time does not depend on the starting point of the retention interval. Rate of forgetting must therefore be constant, and it should be possible to superimpose functions that cover different time periods. The negative exponential is the only function that satisfies this constant rate requirement. Empirically, it does well in fitting a wide range of data (Loftus 1985; White 1985; 1991) although there are instances when a power function seems to do better (Wixted & Ebbesen 1991). Exponential forgetting suits a correspondence view because the decreasing probability that discrete elements of information will correspond to the increasingly distant event, and the increasing probability that interfering events will occur in the retention interval, are exponential in nature.

An additional consideration argues in favor of exponential forgetting. Correspondence can be measured by a generalization gradient that is sensitive to the graded nature of accuracy ("dimensional accuracy," sect. 4.2.2.1). If we assume that accuracy decreases with increasing temporal distance in the same way as for other dimensions, the fundamental function relating accuracy to distance is exponential. This relation was proposed by Shepard (1987) to describe the reduction in discrimination with increasing distance along the dimensions of a wide range of sensory continua (the exponential law of generalization). There is no reason to believe that it should not also apply in the case of temporal distance, thus admitting memory to the realm of perceptual phenomena.

It is not difficult to find studies that support the ubiquity of exponential forgetting. A more interesting issue, however, is what it means in the context of the correspondence metaphor. Forgetting functions offer a quantification of remembering, and accordingly are seen to describe changes in the amount remembered. When placed in the context of a correspondence approach, it is inappropriate to translate the function as representing the "decay" of some process. Trace decay or memory loss may be well suited to an account of forgetting in terms of the storage metaphor, but do not fit a correspondence view in which accuracy, and not amount, is the variable of interest. Exponential forgetting must be interpreted in terms of the changes in perceptual or discriminative processes that occur at the time of retrieval.

The second implication of TDR is that remembering at one time is independent of factors that may influence remembering at earlier times. This possibility of temporal independence has not been explored for human memory, although it has been demonstrated for animal memory (White & Cooney, in press). It suggests that independent discriminations are performed at different retention intervals. In studies of children's event memory, for example, correspondences have been sought for memories of an event at different retention intervals (Pipe & Wilson 1994). It will be of interest to ask whether correspondence observed at one time is influenced by factors such as retrieval-cue availability that may improve accuracy at earlier times. If memory and perception follow similar correspondence principles, it should be possible to demonstrate temporal independence.

Approaches that provide an account of memory in terms of encoding processes would not be consistent with demonstrations of temporal independence, because different performances may be evident at different retention intervals even though the same encoding operations apply. Consider the classic study reported by Peterson and Peterson (1959). Recall probability is reduced when a competing task is interpolated in the retention interval. Whatever the storage mechanism, recall is thought to be a function of the amount stored (Laming 1992). What might be the result when

the competing task is terminated partway through the retention interval in the Peterson's procedure? If recall depends on encoding or on how much remains stored at the time the competing task is terminated, further deterioration in performance will, at best, be avoided. Direct remembering and similar approaches that emphasize the correspondence of performance to earlier events allow temporal independence. When the competing task is terminated partway through the retention interval in the Peterson's procedure, a substantial improvement in accuracy should result, thus reversing the direction of the forgetting function and generating nonmonotonicity. This counterintuitive result has yet to be reported, but would contribute convincing evidence that a correspondence view is capable of generating novel predictions.

## Contexts and functions of retrieval

Eugene Winograd

*Department of Psychology, Emory University, Atlanta, GA 30322.*

[genew@fsl.psy.emory.edu](mailto:genew@fsl.psy.emory.edu)

**Abstract:** Koriat & Goldsmith provide an excellent analysis of the flexibility of retrieval processes and how they are situationally dependent. I agree with their emphasis on functional considerations and argue that the traditional laboratory experiment motivates the subject to be accurate. However, I disagree with their strong claim that the quantity-accuracy distinction implies an essential discontinuity between traditional and naturalistic approaches to the study of memory.

I have comments on two issues raised by Koriat & Goldsmith (K&G). The more important of these issues is their expanded conception of retrieval. To put the study of retrieval in historical perspective, "verbal learning" paid little attention to retrieval. Tulving & Pearlstone's (1966) distinction between availability and accessibility made the role of retrieval cues explicit. With the encoding specificity principle, Tulving (1983) focused on the importance of encoding/retrieval interactions. In particular, he emphasized that all conclusions about encoding are constrained by retrieval conditions. This interdependence found further emphasis in transfer-appropriate processing (Morris et al. 1977) and the transfer-appropriate procedures approach (Roediger 1990). Currently, retrieval factors hold center stage, as shown by the enormous interest in explicit versus implicit memory - a retrieval distinction. Now K&G have shown how rewarding the consideration of retrieval in depth can be and what an enormous task awaits us. In their analysis, K&G produce the following list of questions concerning retrieval in any situation; (1) Is the subject being scored for quantity or accuracy? (2) Is the retrieval format recall (production) or recognition (selection)? (3) Is retrieval free or forced? (4) How effective is the subject's memory monitoring? (5) What is the purpose of this act of remembering?

The last question, about the function served by remembering, becomes crucial when accuracy is considered and requires amplification. Michael Ross (1989) has suggested that "people can choose to engage in relatively effortless, theory-guided recall or a more effortful and extensive memory search (p. 355)." Presumably, they would engage in a more effortful search when accuracy is important; remembering in other circumstances might be more reconstructive or inferential. (Choosing one mode or another would be under "strategic control," in K&G's terms). It should be obvious that memory serves social functions and that different settings demand different types of retrieval. When testifying under oath in court, people presumably engage in effortful retrieval, whereas in a social setting such as a dinner party, impression management is an important goal and one may be more concerned with being amusing than with fidelity. Consider what people tell each other about themselves on first dates. Neisser (1988) has made a similar distinction between utility and verity. Utility involves using the past to accomplish some end, as in making a favorable impression; verity means using memory to give

an accurate description. Neisser argues that any act of remembering falls somewhere on a continuum that has utility at one extreme and verity at the other. The assumption that Ross, Neisser, and others share is that remembering can have accuracy as its goal but that it does so less frequently than we professional students of memory think.

It is important to recognize that although social settings vary widely our attitudes as researchers have been strongly shaped by one particular social setting, that of the memory laboratory. The social structure of the memory experiment demands accuracy from the subject. In learning environments - and the first stage of almost every memory experiment requires learning - success is achieved by mastery, not improvisation. Also, it is clear to the subject that the experimenter composed the input and will compare output and input. Furthermore, the short retention intervals commonly used and the fact that the material is usually arbitrary - not "about" anything, as K&G point out - make accuracy appropriate. These are hardly universal conditions, although they do exist in everyday life, particularly in school. For these reasons, it is an exceptional outcome when memory is not accurate in traditional list-learning experiments (see Roediger & McDermott, 1995, for an interesting case of inaccuracy). One salutary result of widening our focus to include both quantity and accuracy is that considerations of function, what memory is "for" (Baddeley 1988), become inevitable. To their credit, K&G perceive this and point us in a sociopsychological direction.

The second issue I wish to comment on concerns my misgivings about mapping the quantity versus accuracy distinction onto laboratory and everyday memory, respectively. In my opinion, the line between quantity and accuracy does not bisect these two methodologies. As they point out, K&G show the effects of their retrieval manipulations in an item-learning format. Clearly, you need not be an everyday memory researcher to study accuracy, although there is an affinity there. More to the point, if one constructs a 2 X 2 table with "quantity" and "accuracy" in the rows and "laboratory" and "everyday memory" in the columns (see Table 1), it is very easy to find studies that fall into all four cells, not just the positive diagonal, as K&G's analysis implies. Here are some examples of the many experiments concerned with quantity using everyday memory procedures (for the lower-left cell in Table 1): Bahrck's study of retention of names and faces of high-school classmates and of former students (Bahrck et al. 1975; Bahrck 1984); diary studies of ordinary life events (Brewer 1988; Wagenaar 1986); and Conway et al.'s (1991) study of long-term retention of knowledge acquired in a course on cognitive psychology. Similarly, there are many studies of accuracy conducted in laboratory settings (upper-right cell in Table 1), as noted by K&G, for example, experiments on eyewitness memory, story memory, and scene or spatial memory.

In short, there is no necessary discontinuity between traditional

and everyday memory as long as both types of researchers keep in mind that their findings are constrained by the social settings in which they occur. In any case, whether one sees continuity or discontinuity between laboratory and naturalistic studies of memory seems to me of comparatively little importance compared to the progress in increasing our understanding of retrieval reflected in the work of Koriat & Goldsmith.

### Hypothesis testing in experimental and naturalistic memory research

Daniel B. Wright

Psychology Department, City University, London, EC1V 0HB, United Kingdom, d.b.wright@city.ac.uk

**Abstract:** Koriat & Goldsmith's distinction between the correspondence and storehouse metaphors is valuable for both memory theory and methodology. It is questionable, however, whether this distinction underlies the heated debate about so called "everyday memory" research. The distinction between experimental and naturalistic methodologies better characterizes this debate. I compare these distinctions and discuss how the methodological distinction, between experimental and naturalistic designs, could give rise to different theoretical approaches.

Koriat & Goldsmith (K&G) were courageous to try to integrate the diverse set of studies, commonly labelled naturalistic, everyday and/or applied memory research, with a single underlying metaphor. How good is the fit between the correspondence metaphor and so-called everyday memory research? K&G admit it "is more a correlation than a perfect mapping" (sect. 2.2, para. 13). The link appears to be based on the view that memories are about events (Conway 1993), and therefore the events, as well as the recollections, should be of interest. However, this theoretical stance does not appear necessary for everyday research, nor does using everyday research require the correspondence metaphor. Although K&G may have identified a good way of classifying some memory research, their division is different from the actual battle lines drawn in the everyday memory debate.

In this commentary I discuss an alternative way of distinguishing types of memory research: naturalistic versus experimental memory research. These roughly map onto Cronbach's correlational and experimental disciplines, which he described in his 1957 APA Presidential address. Naturalistic memory research is where, for a variety of reasons, the researcher cannot assign subjects to groups at random. The brunt of Banaji and Crowder's (1989) attack was on this type of research. Much of what is called everyday memory research uses experimental procedures and therefore does not come under the naturalistic heading, but this work also was largely spared in Banaji and Crowder's attack.

In experiments, researchers randomly allocate people into conditions and can therefore assume that, before the experimental manipulation, these groups are about the same, allowing causal hypotheses to be evaluated - the *ceteris paribus* conditional being satisfied within a known sampling error. If differences on the outcome measures are large enough, researchers attribute causal properties to the differences in the ways the conditions were treated. This is the foundation of much psychological research. One disadvantage of this approach is that in order to assure that "all other things" are basically equal, the conditions that are used are seldom representative of all the situations that the researchers would like their findings to generalize to. This can make it difficult to estimate the magnitude of any effects.

Causal attributions are more difficult to make with naturalistic research. The *ceteris paribus* conditional cannot usually be assumed. The types of hypotheses that can be examined are therefore quite different. Consider an early example of the everyday memory movement that used naturalistic methods: Brown and Kulik's (1977) finding that, within their sample, blacks had clearer memories of various civil rights events than did whites. They did

Table 1 (Winograd). Possible relations between research setting and focus

Research context	Memory measure	
	Quantity	Accuracy
Laboratory	K&G	
Everyday		K&G

Note: Koriat & Goldsmith's analysis suggests that research should cluster in the positive diagonal, as indicated by "K&G" in the cells of this table. There are enough studies in the remaining cells, as indicated in the text, to cast doubt on their argument.

not claim that race caused the difference in memory qualities (according to some philosophies of causation, characteristics like race should never be attributed causal properties, because they cannot be manipulated; Holland 1986), but that race and memory clarity were associated for these events. Association hypotheses are used to estimate the prevalence of certain phenomena and to see how certain characteristics covary in nature. If Brown and Kulik had used better sampling procedures, they would have been able to make claims about the strength of the association between race and memory clarity for these events, but not about why this association occurs.

Whereas K&G's distinction is "metatheoretical," the distinction between naturalistic and experimental is "metamethodological." However, because the types of hypotheses that can be evaluated are different, naturalistic research tends to be more descriptive, while experimental research focuses on process and causal relationships. The use of naturalistic designs may hence lead to particular theoretical approaches (rather than vice versa, as K&G contend). While it could be argued that researchers choose their methodology with a theoretical framework already in mind, much of naturalistic memory research (for example, infantile amnesia, flashbulb memories, very long-term memory for school material, memory for real crimes, etc.) is difficult to investigate with experimental procedures. Hence the choice of topic often determines the design.

Finally, Koriat and Goldsmith (sects. 6.1 and 6.2) argue that the correspondence and storehouse metaphors are distinct, but that each should be used within "methodological pluralism." This applies equally to the distinction between naturalistic and experimental memory research designs. By exploring both the causal mechanisms (with experimental designs) and their impact (with naturalistic designs), researchers can offer a more encompassing science. The differences between these approaches should be recognized so that we can take advantage of their complementary nature.

#### ACKNOWLEDGMENTS

This manuscript was prepared with the support of a British Academy Fellowship. I thank John Gardiner for helpful comments on an earlier draft.

## Authors' Response

### The correspondence metaphor of memory: Right, wrong, or useful?

Asher Koriat and Morris Goldsmith

*Department of Psychology, University of Haifa, Haifa, Israel.*  
 akoriat@psy.haifa.ac.il; mgold@psy.haifa.ac.il

**Abstract:** Our response to the commentators covers four general issues: (1) How useful is our proposed conceptualization of the real-life/laboratory controversy in terms of the contrast between the correspondence and storehouse metaphors? (2) What is the relationship between these two metaphors? (3) What are the unique implications of the correspondence metaphor for memory assessment and theory? (4) What are the nature and role of memory metaphors in memory research? We stress that the correspondence metaphor can be usefully exploited independent of the real-life/laboratory controversy, but that a variety of other metaphors, including the storehouse, should also be utilized in order to more fully capture the myriad facets and functions of memory in everyday life.

We thank the commentators for their thoughtful and stimulating responses to our target article. In the article, we

attempted to integrate a rather broad and complex array of interrelated issues at various levels of analysis: meta-theoretical, theoretical, methodological, and empirical. Hence it is not surprising that different commentators chose to address different subsets of those issues from a variety of viewpoints. Our response is organized around several general themes that emerge from the commentaries.

### R1. Motivation and goals of the article

We first address some of the major points concerning the take-home message of the article. Because some of its motivations and goals may have been misconstrued by some of the commentators (and may likewise be misconstrued by other readers), we begin by outlining the essential thread of our argument: (1) Our examination of discussions of the real-life/laboratory controversy revealed three different dimensions around which the controversy may be seen to revolve; these dimensions, although correlated in the reality of memory research, are not logically interdependent. (2) At the same time, a survey of the work carried out under the banner of everyday, ecological memory reveals a unique preoccupation with the accuracy and faithfulness of memory. This preoccupation has little parallel in the traditional, laboratory approach to memory, which has focused almost exclusively on memory quantity. (3) We proposed that the focus on memory accuracy discloses a way of thinking about memory, embodied by the correspondence metaphor, that is different from the one reflected by the storehouse metaphor that has guided traditional laboratory research. (4) Because the meta-theoretical shift toward the correspondence metaphor has not been generally acknowledged, the study of memory correspondence continues to be constrained by theories and assessment methods, originally derived from the storehouse approach, that are not well suited to express the unique concerns raised in many discussions of memory accuracy and distortion. (5) We accordingly undertook to explicate the logic of the correspondence metaphor and to show how its exploitation in memory research and assessment could engender a bona fide psychology of memory correspondence to complement the quantity-oriented tradition. (6) We demonstrated how such an endeavor might be particularly useful in capturing some of the dynamics of memory in real-life situations and at the same time applicable in laboratory research contexts.

**R1.1. Regarding the real-life/laboratory controversy.** As should be clear from the foregoing outline, although our work was prompted by the real-life/laboratory controversy, our main goal was not to explain or resolve the controversy, but rather to explicate the metaphorical contrast that emerges from it and to show how that contrast can be utilized independent of the controversy. Thus, for instance, **Bruce** is mistaken in stating that "the principal issue that it [the target article] attempts to sort out is the difference between laboratory and naturalistic memory research." Because Bruce feels that the everyday/laboratory controversy has essentially dissipated, he believes that our analysis is therefore a "post mortem." A similar concern is expressed by **Kvavilashvili & Ellis**, who state - we hope rhetorically - that "the primary aim of the current target article is to demonstrate that the controversy has not been resolved."

Of course, some other commentators were simply not convinced that our analysis in terms of metaphors captures the essence of the controversy (see sect. R2).

As just noted, however, our primary aim was to use the real-life/laboratory controversy as a vehicle for revealing the more fundamental distinction between the two alternative conceptions of memory. Therefore, whether or not our analysis helps clarify some aspects of the controversy, we entirely agree with **Kvavilashvili & Ellis** that the contrast between the correspondence and storehouse metaphors is "sufficiently important to stand alone without reference to the everyday/laboratory controversy" (see also **Kruglanski** and **Winograd** for similar comments).

**R1.2. Regarding the correspondence/storehouse distinction.** A second misperception of our intention may underlie some of the commentators' attempts to defend the storehouse-guided, quantity-oriented approach to memory against the perceived threat of correspondence hegemony. Several commentators argue that not only is the storehouse metaphor alive and well, but that it can still make valuable contributions to the understanding of memory. In fact, some went as far as to propose a division of labor between the two metaphors: **Bahrick** maintains that whereas the correspondence metaphor is useful for capturing reconstructive memory processing underlying memory distortion, the storehouse metaphor is useful for capturing replicative processing underlying memory loss. **Schwartz** believes that the correspondence metaphor is well suited for inspiring functional models of memory, but that the storehouse metaphor is more suited to guide the development of structural process models (see also **Kruglanski** for a similar view).

We certainly agree that the storehouse metaphor is useful in capturing aspects of memory to which the correspondence metaphor is not suited, although we do not think that the correspondence metaphor need be limited either to reconstructive processing or to functional modelling (for an example of a correspondence-oriented process model that may be applied to "replicative" memory, see Koriat & Goldsmith, in press b). Perhaps, despite our explicit endorsement of "metaphorical pluralism" (see sects. 6.2 and R6.1), we are to blame for the impression that we were trying to bury the storehouse metaphor. First, we deliberately presented an extreme and hence somewhat primitive version of that metaphor in order to more clearly reveal its underlying logic. Of course, we also pointed out that modern storehouse-guided models have evolved considerably in sophistication and in their ability to deal with issues such as representation and retrieval (sects. 2.1 and R5). However, we admit that we were not particularly concerned about demonstrating the viability of the storehouse metaphor. Thus, **Bjork & Wickens** and **McNamara** are quite right in emphasizing that modern quantity-oriented models have come a long way from their verbal-learning ancestors and have much to contribute.

Second, our main interest is in promoting the correspondence metaphor as a viable alternative to the storehouse metaphor in guiding memory research. As we pointed out (sect. 3), unlike the situation in traditional, storehouse-guided, quantity-oriented research, little effort has been invested in explicating the underlying logic of correspondence-oriented memory research and assessment. Thus, an important goal of the target article is to provide a first step

toward filling that gap. This, then, should explain our asymmetric treatment of the two metaphors, which was also noted, rather approvingly, by **Ben-Ze'ev**: "K&G characterize themselves as metaphorical pluralists, but their sympathy to the alternative metaphor is obvious - and I believe rightly so."

**R1.3. Regarding various types of memories.** A third basic objection among some commentators is that the correspondence metaphor is not a suitable metaphor for all types of memory phenomena. For instance, both **Alterman** and **Karn & Zelinsky** argue that the correspondence metaphor is not well suited to capture memory phenomena that fall under the rubric of implicit or procedural memory (Schacter 1987; Tulving 1985). In a similar vein, **McNamara** complains that our taxonomy left out priming phenomena and measures of processing time.

We do not believe that one or two metaphors can possibly capture the entire repertoire of memory phenomena and processes. Clearly, many phenomena fall outside the "focus of convenience" (sect. 6.1) of both the storehouse and correspondence metaphors. What we have tried to do is carve out that aspect of memory involving the explicit recollection of past states and events and to show how the study of such phenomena can be enhanced by an explication of the correspondence metaphor. A large amount of both everyday and traditional laboratory research is certainly devoted to the study of explicit episodic and semantic memories. Thus, although it is worth considering how other aspects of memory might be conceptualized (see sect. R7), we do not think that their omission detracts from the value of our proposal.

## **R2. Value of the correspondence/storehouse distinction for understanding the real-life/laboratory controversy**

As pointed out in the target article (sect. 1), most previous discussions of the everyday/laboratory controversy have revolved around three dimensions of the controversy, what we called the "what" (substantive questions), "where" (context of inquiry), and "how" (methodology) issues. In our analysis, we attempted to show that beyond (or perhaps beneath) these issues lies a more fundamental distinction between two different ways of thinking about memory — the correspondence and storehouse conceptions.

The reactions to this analysis were mixed. On the one hand, many of the commentators indicated, either explicitly or implicitly, that they found the analysis in terms of metaphors useful for their conception of the controversy. Thus, **Fisher** states that our analysis "elevates the laboratory-everyday memory debate to a higher plane than we have seen in recent years." **Larsen** also believes that our discussion of memory metaphors "is timely and potentially very useful, considering that the controversy over naturalistic versus laboratory approaches has partly stymied theoretical development in this area for more than a decade." **Newby & Ross** maintain that our analysis "offers the promise of theoretical and empirical advancement, as well as a rapprochement between the two traditions." Finally, **Neisser** believes that the distinction in terms of metaphors "makes it possible to see the dispute between 'ecological' and 'traditional' approaches to memory in a new and clearer light." Neisser, however, believes that a somewhat different

conceptual distinction - between a storehouse and an "action" metaphor - underlies the traditional and everyday approaches, respectively (see sect. R7).

On the other hand, several other commentators expressed reservations about the applicability of our analysis to the real-life/laboratory controversy. First, as mentioned earlier, some (**Bruce, Kvavilashvili & Ellis**) believe that the controversy is over and there is no sense in reviving it. They point out that the prevailing state of memory research indicates "a far broader mix of problems, methods, variables, and theoretical orientations than the comparatively narrow study of memory that marked the first one hundred or so years of our science" (Bruce), and hence, "any tension between the two approaches is being gradually resolved in favour of peaceful coexistence and mutual benefit" (Kvavilashvili & Ellis). We certainly hope that this is true. Although we see no harm in "provoking a fresh debate" (Kvavilashvili & Ellis) along what we think are clearer and more fundamental lines, as mentioned before, our primary purpose is to learn from the debate, not to revive it or to explain it.

Second, some commentators have preferred to maintain allegiance to one or more of the "what," "where," and "how" dimensions. **Wright**, for instance, claims that it is the "how" dimension that is at the core of the controversy (cf. Banaji & Crowder 1989), construing this dimension in terms of the contrast between experimental and correlational methodologies. According to him, the choice of methodology ("how") is often dictated by the "where" - laboratory versus real life. However, his characterization of naturalistic research as one in which the experimenter "cannot assign subjects to groups at random" would probably offend most proponents of naturalistic memory research, who argue their case out of strength, not out of weakness. **Bruce**, in contrast, asserts that, if anything, only the "what" issue is germane to the conflict, and that the "what" is motivated by the "where." Thus, he claims that "recognizing the natural contexts of memory would inevitably lead to a broader array of questions and a wider recognition of significant variables. That was the issue, nothing more, nothing less." It is curious that the "where" dimension, which is typically used to label the controversy, is the one that has received the least emphasis in the commentaries. In fact, **Bjork & Wickens** argue that the laboratory/real-world aspect is simply irrelevant, as can be learned from advancements in other sciences (see also **Kvavilashvili & Ellis**). Finally, **Neisser** emphasizes that neither the "where" nor the "how" are essential to the controversy, stating that "although differences of method are often involved, the most fundamental difference - e.g., between my views and those of Banaji and Crowder (1989) - does not concern how research should be controlled or where it should be conducted; it concerns how we think about memory itself."

We believe that the foregoing remarks largely confirm the state of affairs that we described in the target article. There is little agreement about which of the dimensions of the controversy - the what, where, or how - is the most critical, or indeed whether any of them is critical. It is this situation that led us, among other things, to seek a more fundamental distinction in terms of the underlying conception of memory.

A third type of objection, however, concerns the adequacy of our proposed mapping between the correspondence-storehouse distinction and the everyday-laboratory ap-

proaches. Despite the positive comments by many of the commentators, others were concerned that the relationship may be too imperfect to be of real value. This point was made most directly by **Winograd**. He states that "clearly, you need not be an everyday memory researcher to study accuracy, although there is an affinity there." He points out that many studies of memory accuracy are conducted in the laboratory, and that many studies focusing on memory quantity are conducted in naturalistic settings. Similarly, **Kruglanski** argues that our remarks regarding the lack of necessary interdependence between the what, where, and how aspects of the controversy should obtain for the relationship with the metaphors as well.

Perhaps the simplest way to argue the case for the correlation between the real-life/laboratory distinction and the accuracy/quantity distinction is to walk the reader through the same trail that we took. The original impetus for our work (see Koriat & Goldsmith 1994) derived from an apparent inconsistency between the findings from a naturalistic study reported by Neisser (1988b) and traditional laboratory findings. Upon further examination, we found the inconsistency to implicate, among other things, a tendency for the two types of research to focus on different memory properties — accuracy and quantity, respectively. This tendency is easy to see: on the one hand, the heightened concern with issues of memory accuracy and distortion in everyday memory research can be illustrated by looking at almost any edited book or conference proceedings. Consider, for instance, several representative titles from the program of the SARMAC conference held in Vancouver in July 1995: "Accuracy and distortion in the recall of autobiographical memory content," "Stability and accuracy of self-perceived memory change: A longitudinal analysis," "False childhood memories: Research applications and theory," "Confidence and accuracy in eyewitness studies: Is the conclusion changing?" On the other hand, leafing through some of the traditional memory textbooks (e.g., Crowder 1976, Gregg 1986), as far as we could determine, the words "accuracy" and "distortion" are not even mentioned!

In our mind, this relationship could not be a mere accident. **Fisher** expresses the idea nicely when he reiterates our belief that "there is a nonarbitrary link between research in everyday memory and the use of an accuracy-oriented approach, that is, there is something inherently compatible about the marriage between everyday memory and the correspondence metaphor." In this regard, we think that perhaps **Neisser** is being more prescriptive than descriptive when he asserts that "doing," rather than correspondence, is the metaphor underlying the everyday memory approach (compare his concluding remarks in Neisser 1988b).

Note that the relationship between the metaphors and the approaches is stronger than **Winograd's** analysis implies, if we replace his focus on naturalistic and laboratory research contexts with our focus on the everyday and traditional research approaches, respectively. It is perhaps unfortunate that the labels applied to the two camps (particularly the term "laboratory") tend to focus attention exclusively on the context of inquiry dimension. We think most would agree that Elizabeth Loftus' s (1979a) work on eyewitness testimony, for instance, or Marcia Johnsons (e.g., Johnson et al. 1993) work on reality monitoring (cf. comment by Conway) are prime examples of "everyday

memory" research, even though such work is carried out within the four walls of a laboratory.

Of course, as we acknowledged (sect. 2.2), the correlation between the metaphors/properties and the everyday and traditional approaches is not perfect. Clearly, output-bound accuracy measures have sometimes been used in traditional memory research (see **Bjork & Wickens**), and quantity-based measures of memory are often used in everyday memory research. However, it is hard to deny that the study of everyday memory phenomena has brought with it an unprecedented interest in memory accuracy compared to the very limited role that accuracy has played in traditional laboratory-based research and theorizing (**Conway** goes as far as to assert that our argument "that laboratory research is essentially concerned with counting [memory] traces . . . is undeniably correct"). Essentially, the attributes of the correspondence metaphor represent our attempt to synthesize what is common to a great deal of accuracy-oriented everyday memory research.

### R3. Implications of the correspondence metaphor for memory assessment

As mentioned earlier, one of our major aims in explicating the correspondence metaphor was to clarify its unique logic and its implications for the study and assessment of memory performance. Many of the comments acknowledged the value of this metaphor in bringing to the fore aspects of memory that are not well captured by the storehouse metaphor. **Kruglanski**, for instance, states that the correspondence metaphor focuses research attention on issues that are of "paramount relevance to everyday memory concerns," and affords in addition the development of useful new methodologies for memory research." Both **Bahrick** and **Fisher** emphasize the value of the metaphor for the analytic assessment of memory. **Mazzoni** and **Schwartz** stress the contribution of the metaphor in highlighting the role of metamemory processes. However, these and other commentators also bring out several important issues that can help clarify various facets of the assessment of memory correspondence.

**R3.1. What is the proper criterion?** Perhaps the most basic issue concerns the criterion for assessing correspondence, or, as **Begg** puts it, "what corresponds to what?" **Begg** points out that, in most real-life situations, we have no way of knowing what really happened, that is, no objective criterion against which the memory report can be assessed (see also **Newby & Ross**). Moreover, many commentators wonder whether in principle memories should be compared to an "objective" criterion defined in terms of external reality (**Algom, Begg, Conway, Kruglanski, Mazzoni, Newby & Ross, Palmer**). They argue that perhaps it is the rememberers initial perception or encoded representation of the actual event that should constitute the proper criterion. The argument made by **Newby & Ross** is representative:

Individuals may experience the same event quite differently. . . . Consequently, it is not clear what the test of correspondence should be. Perhaps researchers should evaluate memory against an individual's initial representation of the event, rather than against the supposed objective stimulus. After all, we cannot ask more of memory than that recollections reflect the person's original reality; otherwise, we confuse differences in memory with differences in perception."

Along similar lines, **Conway** distinguishes between accuracy and veridicality: "a memory might be completely accurate in that it corresponds directly and fully to some knowledge structure," though it does not follow that it is veridical, and **Algom** points out that "it is perception alone that substantiates reality."

These remarks pose some challenging issues for correspondence-oriented memory research. It is important to note, however, that these issues are not unique to the assessment of memory correspondence, and in fact pertain to the quantity-based assessment of memory as well: Can one calculate "percent recall" for a free-recall task if the list of items actually presented to the subject is unknown? What if the failure to recall a particular item is due to deficient perceptual processing rather than to deficient "memory?" Conversely, is it possible that the correct recall of a particular item actually constitutes a commission error (i.e., is an "adventitious outcome"; **Palmer**)?

In fact, it is rather conspicuous that such questions are hardly ever raised in the context of traditional, quantity-based memory assessment. Perhaps this is because the problem is circumvented somewhat by the typical presentation of sterile and unambiguous stimulus materials that are devoid of personal meaning under tightly controlled conditions (wasn't that **Ebbinghaus's** ultimate aspiration?). Hence, an objective description of the input may be assumed to approximate what was actually encoded. The situation is more complicated when it comes to meaningful (**Conway**) naturalistic memory situations, which allow much more room for idiosyncratic variance in the initial encoding of the event. Thus, it would seem that the issue does not so much implicate the correspondence-storehouse distinction as it does the distinction between naturalistic and laboratory research contexts: The potential discrepancy between subjective and objective memory criteria should pose a greater problem in naturalistic research settings regardless of whether a correspondence or a storehouse metaphor is adopted. Even here, though, lacking direct access to subjects' initial representations, the simplest (and certainly most common) way of handling the problem might still be to use the "objective" criterion as the best estimate of the subjects' initial encodings (but see **Ross**, in press, for some further suggestions).

We should stress, however, that the criterion issue is not just methodological, but metatheoretical, and in fact resembles the one we noted with regard to the treatment of metamemory in memory assessment (see sect. 5.3.2): Should the initial encoding process be considered as part of memory itself, or rather as something that should be controlled for or partialled out in the attempt to assess "true" memory correspondence? If, as is often the case, encoding processes are considered to be an integral part of memory (e.g., **Craik & Lockhart 1972; Tulving 1983**), then perhaps the "objective" description of the event should in any case be the proper criterion. Of course, it would still be useful to have some way of separating the contributions of the initial encoding and those of the other memory processes. Perhaps some method could be devised like the one we proposed for separating metamemory and retention. Clearly, however, this will be no easy task. In this respect, we take the liberty of adapting **Palmer's** comment somewhat: "As jurors we would care only about correspondence with an external state of affairs, but as students of memory, we care about" *both* the correspondence with an external

state of affairs, *and* "the relationship between past experience and current behavior."

**R3.2. The wholistic assessment of correspondence.** In section 4.1 we outlined the type of wholistic memory measures that follow uniquely from the correspondence metaphor (**Nelson** objects to the term "unique" in this regard; perhaps we should have said "most naturally"). These measures capture the overall multidimensional fit between a complex memory report and some objective description of an earlier event. This is in contrast to the typical focus on analytic, item-based assessment methods that dominate memory research in both laboratory and naturalistic contexts. Our aim in discussing these measures was not only to promote their development and use, but also to explicate some of their special features, in particular, the fact that they cannot be implemented independent of functional considerations. Indeed, several commentators stressed the functional view of memory even more than we did (see sect. R7). In addition, we emphasized that such wholistic measures must also be tailored to specific memory domains and tasks. **Neisser** provides a good illustration of this point, noting that the global measure he used in one study dealing with hearing the news about the *Challenger* disaster (Neisser & Harsch 1992) had to be modified for use in another study dealing with hearing about an earthquake (Neisser et al., in press).

**Neisser's** example, together with the various criterion issues considered earlier, reinforces our discussion of the problems involved in deriving wholistic correspondence measures, and in fact Neisser observes that "in some cases the assessment of correspondence and accuracy is so difficult that no resolution is possible." We hope, however, that these difficulties will not deter researchers from accepting the challenge of developing such measures. Indeed, we were encouraged by **Kruglanski's** endorsement of our belief in the potential benefits of importing novel measurement techniques from the domain of social perception.

**R3.3. Evaluating correspondence on continuous dimensions.** Most contemporary memory research is conducted using what we called "analytic" assessment procedures. Among these, the evaluation of dimensional accuracy provides an interesting case: Although it is much more common in both laboratory and naturalistic research than the wholistic approach, researchers generally fail to realize that it too reflects a correspondence rather than a storehouse metaphor.

The implicit use of a correspondence metaphor in dimensional assessment is nicely illustrated by "memory psychophysics" (Algom 1992). Just as traditional psychophysical scaling methods are designed to determine the mapping between physical stimulus dimensions and their perceptual correlates, memory psychophysics investigates the mapping between physical stimulus dimensions and their memory representations. Thus, in the target article (sect. 4.2.2.1), we cited Algom et al. (1985), who found that for the particular dimensions studied (height, width, and area of rectangular stimuli), both memory and perceptual representations mapped onto their physical referents by the same type of functional relation - a power transform. In his commentary, however, **Algom** describes a further example of work that he has done on women's memory of labor pain (Algom & Lubel 1994), which he believes eludes treatment in terms of memory correspondence. Algom

argues that because of the lack of an objective criterion, or even common units for the physical and remembered stimulus (intrauterine pressure and pain, respectively), correspondence is not at issue. However, we think that his use of the term "correspondence" is overly restrictive (which may also explain his reservations regarding the role of correspondence in the study of perception). In our conceptualization, his work deals precisely with the (rather complex) correspondence between memory and the external world, as well as between memory and perception. We agree, however, with his observation that in those cases in which common units are lacking, it is not possible to speak of "truth" in any meaningful sense. (See also **Palmer**, who suggested that behavioral learning principles might provide a "tool kit" for extending the study of memory correspondence to situations where common units are lacking.

Another innovative example is provided by **Bahrick** (Bahrick et al., in press), who derived both an accuracy-based measure of distortion and a quantity-based measure of memory "loss" in studying people's memory of their high school grades. He points out that computing separate measures for item loss and item distortion is analogous to deriving the constant and variable error components by the method of average error in psychophysics. This approach allowed him to compare the time course of forgetting in the sense of both loss and distortion, and to clarify the relationship between the two. **White** also emphasizes the similarity between memory and perception, and discusses ways in which the measurement of dimensional memory accuracy can be enlightened by principles and techniques from the study of perception.

**R3.4. Item-based assessment: Output-bound accuracy versus input-bound quantity.** Turning now to the more standard, item-based approach, the central issues raised by many of the commentators concerned the relationship between quantity-based and accuracy-based memory measures, and the adequacy of output-bound accuracy as a measure of memory correspondence.

(A) Why the distinction? As **Fisher** notes, there is often a practical reason for using output-bound memory measures in real-life situations such as eyewitness testimony, because the experimenter cannot possibly enumerate all of the information making up the input (i.e., the witnessed event). Our point, however, is that output-bound accuracy should be of both theoretical and practical interest, not just by "default," but because it captures a unique property of memory. Note that even without an enumeration of the entire input, it is possible to calculate an output-bound *quantity* score that captures the amount of (correct) information provided by the rememberer, and indeed, this is often done in naturalistic research. Such a measure, however, like the more traditional input-bound, free-report quantity measures (e.g., percent recall), misses an important attribute of the memory report - what we have called its dependability, that is, the extent to which each reported item of information can be depended on to be correct. This attribute, which is of great concern in many situations, is captured by the output-bound accuracy measure.

This is an important point that apparently needs clarification. **Nelson** emphasizes that in computational terms, the "sole difference" between the input-bound quantity measure and the output-bound accuracy measure is that the former includes omission errors in the denominator

whereas the latter does not (see also **Bjork & Wickens, Mayes et al.**)- That, of course, is true. However, conceptually, the distinction between the two types of measures could hardly be more substantial. To illustrate, consider the issue of the credibility of child witnesses, which has gained increasing attention in recent years (e.g., Ceci & Bruck 1993). On the one hand, a fairly robust finding is that children remember less information than adults (e.g., Brown 1979). This finding is certainly relevant if we are concerned that a child witness may not provide as much information as would an adult in a given situation (but see Ornstein et al. 1992). However, it may not be relevant if concern lies in whether or not the court can trust what the child *does* report. That issue relates specifically to output-bound accuracy, which, in contrast to the quantity measure, does not hold the child accountable for what he or she does not assert to be true. Moreover, as we have shown (Koriat & Goldsmith, in press b), the effectiveness of a person's decision to suffer an "omission error" (say "I don't know"), rather than volunteer incorrect information, depends on the effectiveness of his or her monitoring and control processes. Thus, in a sense, the simple computational difference between excluding and including omission errors in the denominator of the memory measure translates into the fundamental theoretical difference between considering and not considering the role of subject control in determining the dependability of the memory report.

(B) Accuracy versus quantity? Of course, this is not to say that output-bound accuracy alone is sufficient to capture the overall quality of a witness' memory or memory report. Several of the commentators were concerned that we had missed the "deep complementarity" (**McNamara**) between quantity and accuracy, pointing out that perfectly accurate testimony (in an output-bound sense) may be virtually worthless if very little information is reported. Thus, McNamara asserts that "at a more global level, accuracy is scaled by quantity: *Ceteris paribus*, who will the jury believe more, . . . the witness who remembers the color of the assailant's shirt but nothing else (low quantity; high [?] accuracy), or the witness who constructs a detailed account of the event, including clothing, setting, the time of day, and so forth (high quantity; high [?] accuracy)?" Likewise, **Bjork & Wickens** took us to imply that omissions are not a serious problem in real-world contexts like witness memory, but argue that "the failure of the witness to remember salient aspects of the criminal episode leads juries to lose confidence in what the witness does report." Similar concerns were expressed by **Kvavilashvili & Ellis**, and by **Begg**, who also reproached us for endowing accuracy with an "aura of virtue" but equating quantity with "bean counting."

Several remarks are in order. First, it is noteworthy that both **McNamara** and **Bjork & Wickens** chose to argue their case from the juror's perspective. What juries believe and why is indeed an interesting and important question (see, e.g., Duggan et al. 1989), but in fact that issue is independent of the one concerning the actual relationship between the amount of information reported by a witness (input-bound quantity) and the dependability of that information (output-bound accuracy). Jurors may believe that there is a positive correlation between these two memory properties (or perhaps they, like some researchers, simply fail to distinguish between them), but whether or not they are right is an empirical question. Certainly quantity and

accuracy measures can be dissociated, as our own work has shown (see also **Bahrick, Fisher**).

Second, despite our admitted bias in focusing on the value of the various types of accuracy measures, we thought we were clear in acknowledging that one of the advantages of the item-based approach is that both the quantity and the accuracy of the reported information can be evaluated in the same procedure (sect. 4.2.2.2), and that they should generally be considered in tandem (see also Klatzky & Erdelyi 1985). Indeed, the quantity-accuracy profile (QAP) methodology that we proposed (sect. 5.3.2; see also Koriat & Goldsmith, in press b), allows one to do just that: "Compared to the standard point measures of memory performance, the derivation of quantity-accuracy profiles allows a more global evaluation of potential memory performance in terms of both accuracy and completeness." However, what the proposed QAP methodology does not do is "scale" accuracy by quantity, or vice versa, in deriving a single composite memory score (as does the signal-detection measure  $d'$  for forced-recognition memory). Indeed, because each memory property is of interest in its own right, it is advantageous to be able to examine accuracy and quantity separately (see **Bahrick**). Moreover, as pointed out in the target article (sects. 4.1 and 5.3.2), if an overall assessment of performance is desired, functional considerations tied to the specific circumstances of the testimony or the particular theoretical interests of the researcher will need to dictate the relative weight given to each of the two measures.

Third, the relationship between quantity and accuracy is complicated even further by the potential for differences in the level of generality or "grain size" of the memory report (see sect. 5.1). We thank **Small** for pointing out some classical sources on this topic (among others): "As **Thucydides** and **Aristotle** implied, the opposition is not simply between the quantity of the memories (storehouse) and their accuracy (correspondence) but within the correspondence metaphor, between what kinds of memories are subject to being remembered with specific details and what kinds are remembered only in a general sense." **Fisher** provides a nice illustration of how this factor may underlie too\* some "experimental anomalies" that are created by looking at either accuracy or quantity in isolation. Having failed to find any effect of retention interval on the accuracy of eyewitness testimony in several studies, he rescored the data to take into account possible differences in grain size, and found that "the responses made after long delays were less precise (coarser grain), although equally correct, than those provided after short delays." Thus, Fisher stresses a troubling implication for eyewitness research: "In order to meaningfully compare response accuracy across two experiments, one needs to ensure that the witnesses in both experiments were similar on the dimensions of report option and grain size." Clearly, more methodological and theoretical work is needed to meet the challenges presented by both report option and control over grain size in the study of memory accuracy.

#### R4. Implications regarding metamemory and memory

In discussing the implications of the correspondence metaphor, we pointed out that subject-controlled metamemory processes play an important role in the strategic regulation

of memory accuracy, particularly in real-life memory situations. Our analysis focused on what **Fisher** correctly points out are "post-ecphoric" processes (Tulving 1983) and leads to what **Winograd** calls "an expanded conception of retrieval" (see also Moscovitch, in press, and Barnes et al., 1995, for similar recent proposals). We tried to show how a more careful consideration of such processes, together with the distinction between memory quantity and memory accuracy, can help resolve some apparent anomalies in the literature and provide new directions for theoretical and methodological development. One such direction is the attempt to distinguish the separate contributions of retention, monitoring, and control to free-report memory performance. There seems to be a general consensus among the commentators about "how rewarding the consideration of retrieval in depth can be and what an enormous task awaits us" (Winograd).

**R4.1. The importance of metacognition.** Our emphasis on the impact that metacognitive processing has on memory performance and its "practical and theoretical importance" is endorsed by **Nelson**, who also points out that metamemory is an important contributor not only in naturalistic memory situations, but in traditional laboratory research as well. No argument there. There is a vast array of metacognitive processes that have been identified and/or studied in the laboratory (see, e.g., Metcalfe & Shimamura 1994; Nelson & Narens 1990; 1994; Schneider & Pressley 1989) and probably many more that are waiting for attention. Indeed, our own research (Koriat 1993; 1995; Koriat & Goldsmith 1994; in press b) has been entirely laboratory-based. Our point is simply that these processes generally operate more freely both in everyday memory situations and in naturalistic research. Therefore, their study is particularly crucial for those interested in understanding the dynamics of real-life remembering.

Indeed, **Fisher** also stresses this point, noting that in contrast to traditional laboratory research, which often does its best to eliminate the contributions of metamemory, in everyday memory research "we often do not have the luxury of eliminating or controlling these nonmemorial factors, and so they become an integral part of the eyewitness recollection process." He, however, considers decision processes such as report option and control over grain size to be "principles of communication and not memory per se." We would be reluctant to adopt the term "communication" for these processes. Although the term is useful in emphasizing their sociopsychological context (**Winograd**), we believe that it misses the intrinsic role that self-directed monitoring and control processes play in determining what one actually believes one remembers. That is, the decisions underlying overt responses, such as "I don't know" or "it happened around six o'clock" (rather than precisely at six), may be made not only for the sake of communicating one's memories to others, but may also constitute covert self-attributions that affect what a person actually "remembers" (see **Begg**).

This idea is brought out nicely by **Newby & Ross**, who draw our attention to an illuminating parallel between the processes of monitoring one's own memory and those involved in monitoring the accuracy of other people's memories. Ross (in press) identified various criteria that people use when judging the validity of their own or other people's memories, including such factors as vividness, presumed

memorability of the event, internal consistency, and consensus - whether other people remember the event in the same way. It is easy to see (Newby & Ross give several examples) how such factors could influence both self- and other-attributions of memory accuracy, and in effect determine what one believes to be true.

**R4.2. Separating memory and metamemory.** Other comments were directed at issues concerning the proposed separation between retention, monitoring, and control. With regard to memory and monitoring, **Mazzoni** brought up some interesting points concerning the need to distinguish between accuracy from the experimenter's viewpoint and accuracy from the subject's viewpoint. She correctly implies that what we have been calling "monitoring effectiveness," that is, the correspondence between one's confidence judgments and the actual correctness of one's answers, is defined from the experimenter's perspective. A person may be highly confident in an incorrect answer, but this may "accurately" reflect the person's memory representation (see Koriat 1995). An important implication of her remarks is that in order to fully understand the effects of poor monitoring from the experimenter's viewpoint (e.g., Koriat & Goldsmith in press b, Experiment 2), experimenters will need to more fully understand the determinants of monitoring and monitoring accuracy from the subject's viewpoint.

**Schwartz** presents a view very similar to ours (though he may not realize this) regarding both when and how one might try to separate between retention, monitoring, and control. He argues that one's treatment of metamemory should depend on whether one is interested in developing functional models or structural/process models of memory. Whereas the development of functional explanations of everyday remembering dictates that metamemory processes be allowed to operate freely, in developing structural or process models it is crucial to distinguish between the various memory and metamemory components, because each can affect memory in different ways. The approach that he proposes to separate the components is essentially the one that we have utilized in our own research — to study one component while holding the others constant (see Koriat & Goldsmith, in press b). Of course, we should emphasize that in the context of such a research strategy, holding metamemory constant (e.g., by using forced-report procedures) is not the same as ignoring the effects of subject-controlled processes (see also **Nelson's** endorsement of the need to make metamemory explicit in the assessment of memory performance).

**R4.3. Methodological problems: Subject control and output-bound accuracy.** Two of the commentaries raised methodological objections that may be seen to involve the relationship between metamemory processes and the distinction between input-bound and output-bound memory measures (see related discussion in sect. R3.4.A). We will consider each in turn. **Bjork & Wickens** argue that because the output-bound accuracy measure is a conditional statistic, based only on those items for which the subject decides to volunteer an answer, it is subject to "the complex and confusing selection artifacts that have always bedeviled such measures." Specifically, they point out that output-bound accuracy might sometimes be "higher for poor study conditions, more difficult materials, and less alert subjects." For instance, a distracted subject might choose to report

only the few items that were so salient (perhaps idiosyncratically) that they couldn't be forgotten, whereas a more alert subject might report many more items, but demonstrate lower output-bound accuracy. Bjork & Wickens ask, can we really say that the former subject's memory is more "accurate?"

**Bjork & Wickens** note that the problem they raise has plagued free-report memory assessment for many years, and indeed, it is this very problem that makes the signal-detection methodology inapplicable to free-report situations (Lockhart & Murdock 1970). Why does item selection pose a problem for researchers employing the signal-detection methodology, and for Bjork & Wickens? Essentially, it is because these researchers (see also **Mayes et al.**, discussed below) are interested in measuring memory "accuracy" in an *input-bound* sense, that is, how well the subject's memory reproduces the entire input list (or event). Allowing the subject to choose which items to answer, and looking only at the correctness of those answers, means that the set of items on which the (output-bound) percentage is based may not be a representative sample of the input. Clearly this won't do for a useful input-bound measure.

However, at the risk of repetition, when the research focus is on output-bound accuracy, then (1) one is interested in the dependability of the information that is reported, rather than in the amount of recovered information, and (2) it is subject control in selecting which items to answer (i.e., the option of free report) that operationally distinguishes the former property from the latter (sect. 4.2.2.2). Thus, the concern with output-bound accuracy implies a concern precisely with the products of subject control, that is, with the selection effects themselves (see sect. R3.4.A).

From this perspective, the examples pointed out by **Bjork & Wickens** illustrate the conceptual distinction between input-bound and output-bound measures of memory performance and their relationship to metamemory processes. Suppose that we were to focus on one particular item of information volunteered by both their alert and distracted subjects. Whose statement should we trust more? Probably that of the distracted subject, even though he or she reported fewer items overall. The fact is, each of the distracted subject's statements is more likely to be correct than each of the alert subject's statements, given the difference in output-bound accuracy.

Indeed, to the extent that such dissociations between quantity and accuracy are reliable, they call for a detailed analysis of the separate contributions of retention, monitoring, and control to memory performance. Under what circumstances will dissociations emerge? Could more difficult tests, poorer viewing conditions, and so forth impair overall retention (or encoding), but at the same time yield a more polarized monitoring distribution (i.e., either you know it, or you don't; either you saw it, or you didn't, etc.)? If so, better monitoring resolution could lead to superior, or at least equivalent, output-bound accuracy despite the poorer retention (see the comparison of recall and recognition performance in Koriat & Goldsmith in press b, Experiment 1). Similarly, as **Bjork & Wickens** imply, the word "pumpkin" might be remembered with high confidence and have a high probability of being volunteered, even under superficial encoding conditions or by an inattentive subject, simply because the subject's metamemory is oper-

ating effectively ("I'm sure I remember 'pumpkin,' because that's what I call my girlfriend").

We agree that there may be better ways of assessing memory for those interested purely in input-bound quantity/accuracy, for instance, eliminating the selection problem by using a forced-report procedure, as **Bjork & Wickens** suggest. However, for those interested in evaluating output-bound accuracy and understanding its underlying mechanisms, that would be like throwing out the baby with the bath water. Moreover, such an approach would exclude some of the important dynamics underlying memory performance in real-life situations, such as eyewitness testimony. In this regard, Bjork & Wickens maintain that "courtroom testimony is seldom free report. Witnesses are rarely permitted to give narrative answers and are often forced to answer questions." We think that this is a bit overstated. Regardless of whether the witness responds in a free-narrative style or answers specific questions (which is a test-format variable), he or she is always allowed to say "I don't know/remember" if he or she actually doesn't know or remember. That is, report option is always "free." Of course, as Bjork & Wickens point out, there are often both implicit and explicit pressures to supply answers in courtroom testimony (as well as in other memory situations), and these may act to lower the witness' response criterion. However, the way in which witnesses will accommodate these and other demands into their control policy will probably depend on social, functional, and metamemorial factors (e.g., possible age differences in subject control; see Moston 1987), all of which we believe deserve further study.

We now turn to the commentary by **Mayes et al.**, who analyze the item-based quantity and accuracy measures that we discuss from the standpoint of signal-detection theory (SDT). We believe that their analysis is somewhat misguided, mainly because it assumes that we too are motivated by the traditional (signal-detection) desire to achieve a single global measure of (input-bound) memory accuracy. Thus, they argue that our accuracy measure, which ignores omissions, is not a good measure of memory correspondence. What is needed, they say, is an overall correspondence measure that takes both omission errors and false alarms into account: "SDT provides such a measure for recognition, but not for recall, as K&G argue. Our contention is that neither does their accuracy measure; nor at present, does any other recall measure."

Of course they are quite right. As discussed earlier (sect. R3.4.B), neither the output-bound accuracy measure nor the input-bound quantity measure alone can fill that job (but see our discussion of wholistic correspondence measures in sect. 4.1). In tandem, however, they do allow the researcher to focus on two important properties of memory in free-report situations: its quantity and the extent to which it can be depended on to be correct. As interesting as **Mayes et al.'s** analysis is, it seems to miss the point that in addition to posing some methodological limitations, free-report memory situations also endow a different *meaning* to many of the concepts used in signal-detection analyses of forced-recognition performance. For instance, it is simply inappropriate to apply free-report accuracy and quantity measures to the old/new (yes/no) recognition paradigm, as Mayes et al. have done, because by doing so, these measures lose their intended interpretations (see note 14 in the target article). Thus, the free-report quantity measure is not

equivalent to the hit rate (which can be arbitrarily raised to any desired level), and the free-report accuracy measure is not equivalent to hits/(hits + false alarms), because this latter proportion does not depend on the subjects actual commitment to the correctness of his or her answers.

Finally, **Mayes et al.**'s use of the term "monitoring" is also different from ours. Although they assert that for us, "monitoring is discrimination between target and foil items," this in fact appears to be their use of the term, and in keeping with the signal-detection approach, they expect that monitoring should provide the basis for good quantity performance. We, however - borrowing from the metacognitive judgment literature — treat monitoring as the subject's ability to discriminate between correct and incorrect answers, that is, which items he/she can answer correctly and which he/she cannot. This distinction is a subtle one, and we thank Mayes et al. for helping us clarify it. As these researchers discuss, monitoring (as they use the term) may be involved in arriving at a best candidate answer for a particular question (i.e., by eliminating the myriad of alternative possible responses), but once that best candidate is chosen, monitoring (as we use the term) also determines one's confidence that the answer is correct and contributes to the decision whether to provide the answer or to abstain. Mayes et al. correctly point out that in recall testing we cannot evaluate subjects' monitoring effectiveness in the first sense. However, we can in the second sense, by computing the correlation between confidence and correctness across items under forced-recall instructions (see Koriat & Goldsmith, in press b).

## R5. Implications for memory theory

Although the implications of the correspondence metaphor are perhaps most salient with regard to the way in which memory is assessed, the focus on memory correspondence should also influence memory theorizing. Indeed, we argued (sect. 3) that many contemporary approaches to memory seem to reflect an implicit shift toward a correspondence-oriented conception. However, the dividing line between the correspondence and storehouse approaches may be quite fuzzy. Thus, for instance, **McNamara** points out that theories such as ACT\* (Anderson 1983), SAM (Gillund & Shiffrin 1984), and TODAM (Murdock 1982) have incorporated much more sophisticated schemes of representation and processing than earlier models. One might wonder, then, whether the metaphorical contrast we proposed is in fact useful in the realm of memory theory. This calls for a point of clarification regarding two concepts - "storage" and "representation." The notion of "storage" in its broadest sense is so basic to our thinking about memory that it is very difficult to eliminate it altogether from our theorizing. Thus, most theories posit either implicitly or explicitly that information must somehow be held between exposure and test. (An interesting exception that proves the rule is **White's** characterization of "direct" memory, in which memory is conceived almost literally as the perception of the past; see also Watkins 1990 and sect. 3 of the target article.) The form in which the information is held, however, is a matter of dispute. **Ben-Ze'ev**, for instance, distinguishes between "storing" and "retaining," the former implied by a "container approach to the mind," and the latter implied by a view of memory as a "disposition" or "capacity" (e.g., for

Correspondence) that has the potential to be actualized. He argues for the latter view. **Kruglanski**, on the other hand, believes that storage constitutes part of the "critical nucleus" of memory, whether or not it is conceived as a "specific and discrete recording of an event," or "diffuse and widespread modifications of the whole cognitive system (Craik 1983)."

The important point to note is that the incorporation of some notion of storage or retention does not imply subscription to the storehouse metaphor, at least not in the sense that we emphasized. What we took as the distinctive feature of a storehouse view is the way in which the "stored" information is treated: as a set of items ("ideas," "images," "records") that can be counted. This characteristic, a hallmark of the influential verbal-learning tradition, allows memory to be evaluated by a simple count of the number of items remembered after a retention interval. Of course, one may conceive of memory as a store of ideas, as Locke did, or as a store of images, and be concerned instead with the extent to which these ideas or images conform to reality. This kind of treatment would then imply a *correspondence* metaphor. It is in this sense that **Ben-Ze'ev** asserts that the correspondence and storehouse conceptions are not mutually exclusive, and we agree. In fact, our own work within the item-based approach is perhaps a good example.

A somewhat similar distinction can be applied to "representation." As **Conway** correctly notes (and see **McNamara**), the concept of representation is one of the cornerstones of the cognitive approach to memory. The virtue of representational models is that they must address some qualitative aspects of memory, notably the content of what is retained. The way in which that content is treated in presenting and testing the model, however, is another matter. The use of a representational model implies a correspondence view only insofar as the memory representations are in fact treated as descriptions of, or as being "about" past events, and the model is then evaluated accordingly. Thus, as McNamara protests, many contemporary theories incorporating sophisticated representational schemes "are founded on and evaluated against data produced by quantity-oriented research, such as learning lists of words or sentences." Indeed. But have these models also been tested to see how well they capture data pertaining to the *correspondence* between people's memory reports and past events?

We believe that making the correspondence metaphor explicit can help lead to theories and models that are specifically correspondence oriented. By this, we mean, among other things, models that are designed to explain both how memory correspondence is achieved and why it can go wrong. We admit that the correspondence metaphor, unlike the storehouse metaphor, does not in itself provide any guidance about what such a theory should look like (but see the discussion of correspondence-type metaphors in sect. R6.3). Thus, a correspondence-oriented model based on a reconstructive view (cf. **Larsen**), might look very different than a model based on a "direct" Gibsonian view of memory (cf. **White**). However, as mentioned earlier, we disagree with **Schwartz's** contention (see also **Kruglanski**) that the correspondence metaphor is suitable for functional models only. The correspondence metaphor can and should also generate structural or process models (and we include here connectionist models) that focus on the memory — and metamemory — mecha-

nisms underlying memory accuracy and distortion (e.g., Koriat & Goldsmith, in press b; Metcalfe 1990; Wagenaar & Boer 1987).

## R6. Regarding the metaphors

We now turn to issues concerning the metaphors themselves. One general conclusion is quite clear from the commentaries: researchers take metaphors seriously! Aside from the many comments directed to the correspondence metaphor, concerning both its advantages and its limitations for guiding the study of memory, there were also advocates (as well as critics) of the storehouse metaphor and of several other proposed metaphors. Two notable exceptions to the general debate are the commentaries of **Bruce** and **McNamara**, both of whom gave arguments to the effect that "memory research moves in directions that are independent of abstract background metaphors" (Bruce).

**R6.1. Metaphorical pluralism.** A general theme that runs throughout the commentaries is the need for metaphorical pluralism (but see **Bruce**, who believes that "reasonable as that call is, it is unlikely to lift the audience to its feet"). As mentioned earlier (sect. R1.2), this view is implied by many of the commentators who argue for the value of both the correspondence and the storehouse metaphors. **Nelson** was most explicit in emphasizing that "investigators should Use whichever metaphors work best. . . and that in contrast to the substantial achievements in philosophy of science for evaluating theories, no satisfactory method is available for evaluating metaphors and analogies." Similarly, **Neisser** also stresses that there is no point in arguing about metaphors "as if an empirical question were at stake." This view, of course, is precisely what we expressed in the target article (sect. 6.2).

There seems to have been some misunderstanding, however, of our final message regarding the "chariot of science." **McNamara**, for instance, states that he is "moved instead by a different metaphor: when it comes to pulling chariots, two horses are better than one" (see also **Kruglanski**). That, however, is what we thought we had said. In addition, **Nelson** expresses his belief that "progress seems to be less likely if two horses pull in different directions - and directions of pulling that are too different may even be counterproductive and pull apart the metaphorical chariot." Perhaps we should have used a different metaphor. Our belief is that science will be best served if each metaphor is exploited to its fullest. Thus, our admonition is to avoid *compromising* the particular advantages of each metaphor, not to avoid constructive cooperation.

**R6.2. Is correspondence a metaphor?** One point that was brought up by several commentators is that correspondence is not actually a metaphor. **Ben-Ze'ev** notes that, unlike the storehouse metaphor, correspondence is not a metaphor about what memory is. **Neisser** states that correspondence is better thought of as a goal or criterion. Finally, **Larsen** asserts that the notion of correspondence "does not confer any surplus meaning from its source domain to the understanding of memory, and it does not suggest any further properties of the memory system. Rather, it identifies a core feature to be included in an alternative metaphor."

We have characterized the correspondence metaphor as an "abstract" memory metaphor. We think that correspondence, as an abstract concept, can be applied to memory in a metaphorical as well as in a literal sense. Take, by analogy, the "activity" metaphor suggested by several commentators (**Alterman, Karn & Zelinsky, Neisser**; see sect. R7). On the one hand, at least some instances of remembering can be characterized as (cognitive) activity in a literal sense. At the same time, however, by considering other kinds of activity, activity as a metaphor (or analogy) may indeed confer surplus meaning from its source domain(s) to the understanding of memory. The same is true for correspondence: the correspondences between a photograph and its subject, a sculpture and its likeness, a percept and its distal stimulus, a model and its referent, a regression line and its data points, all suggest different ways of thinking about memory — what it is designed to achieve, how it might achieve it, and how it might be evaluated. The concept of "goodness of fit" (sects. 2.2 and 4.1), borrowed from statistical analysis, is one such example.

However, there is not much to gain by belaboring this issue. Whether it is in fact a metaphor, an analogy, or simply a conception, our point is that there are important implications that derive from thinking about memory in terms of its correspondence with past events. Moreover, just as there are many possible variants of the storehouse metaphor (e.g., **Bjork & Wickens**), the correspondence metaphor also denotes a class or *type* of memory metaphor, in which (as pointed out by **Larsen**) correspondence is a "core feature" (as well as some of the other interrelated features that we specified in sect. 2.2). These more concrete instantiations will generally provide more specific constraints and guidance for the ensuing research and theorizing, as we now consider.

**R6.3. Correspondence-type metaphors.** A useful illustration of a more concrete correspondence-type metaphor is **Larsen's** "archaeology" metaphor. As he notes, that metaphor assumes a reconstructive approach to memory (Bartlett 1932), and is similar to the "paleontology" metaphor used by **Neisser** (1967). **Larsen** stresses that the archaeology metaphor

suggests a conception of memory that is remarkably close to Koriat and Goldsmith's correspondence view (cf. sect. 2.2) yet does not ignore that traces from the past are indeed stored somehow. To achieve correspondence between a present account and past reality is precisely the overarching aim of archaeology. Like in memory, collecting items from the past only serves the purpose of constructing true descriptions to represent the past.

It is interesting that **Neisser's** current comments point out that memory construction need not serve only the goal of achieving correspondence, so that while the construction metaphor "still makes sense" to him, he apparently no longer endorses a paleontology metaphor. Be that as it may (see further discussion in sect. R7), metaphors such as archaeology or paleontology are indeed good instantiations of a correspondence-type metaphor. Thus, we (and **Larsen**) disagree with **Ben-Ze'ev** when he asserts that the correspondence metaphor is incompatible with the reconstructive approach. His argument seems to be aimed against a passive, "copy" type of correspondence metaphor.

Many other types of correspondence metaphors may also be envisaged (e.g., "resonance," "holography," "stage-setting"; see note 5 of the target article). In fact, the notion

of memory as "perception of the past" is itself a correspondence metaphor, which can be fleshed out further depending on one's particular view of perception. Thus, for instance, the "paleontology" metaphor was proposed by Neisser (1967) as a useful metaphor for capturing both perception and memory. **White**, on the other hand, viewing memory and perception quite differently than Neisser did in 1967, outlines a more "direct" correspondence view in terms of a theory of direct remembering (TDR) (White 1991), following Gibson's (1979) view of perception: "In TDR, memory and perception are continuous and the same discrimination principles apply to both. Environmental information at the time of retrieval allows direct perception of the remembered event. The event is not stored but perceived directly, albeit at a temporal distance." Finally, according to **Kruglanski**, who is perhaps reading in some aspects of social perception, the correspondence metaphor treats memory as a "judgment" about past events, though he argues that a judgment metaphor of memory must be supplemented by some type of storage conception.

In sum, we have chosen to present an abstract correspondence metaphor, rather than a particular version, because our primary concern is in explicating the general logic of the correspondence conception, not in putting forward a specific view of memory correspondence. Clearly, however, this metaphor can submit to a variety of instantiations depending on the researcher's particular metatheoretical convictions.

### R7. Beyond the correspondence metaphor: The myriad facets and functions of memory

Perhaps one of the most salient features of memory is its multitude of facets. Thus, as mentioned earlier (sect. R1.3), several commentators highlighted aspects of memory for which neither the correspondence nor the storehouse metaphors seem well suited, and some proposed their own alternative metaphors to capture these aspects. Common to all of these discussions is some assumption about the basic function of memory in subserving adaptive interaction in everyday life.

Both **Alterman** and **Karn & Zelinsky** emphasize the function of memory in supporting a variety of activities and procedural skills that constitute the major portion of our daily interaction with the environment. For instance, **Alterman** stresses the "pragmatic" aspects of memory in such activities as operating a photocopy machine. Similarly, **Karn & Zelinsky** point out that "memory is most often used, without conscious awareness, in natural tasks such as driving, walking, grasping, speaking, and problem solving," and emphasize "the active role that memory plays in goal-directed behavior." These commentators argue that correspondence or accuracy is not at issue for such activities. This argument would seem to hold for the entire domain of implicit or procedural memory (Schacter 1987; Tulving 1985), in which memory is not really "about" anything (Tulving 1985), and so issues of truth and accuracy are simply inapplicable (can someone accurately or truthfully ride a bicycle, operate a copy machine, or solve an anagram?). Perhaps the correspondence metaphor could be stretched to cover such phenomena, for instance, by considering the correspondence between an organism's current behavior and experienced contingencies in the environment, or between its behavior at one point in time and

another (**Palmer**). It might be more fruitful, however, simply to seek another metaphor that is better suited to capture such phenomena. **Eichenbaum** reaches a similar conclusion based on the parallel between implicit versus explicit memory in humans and "stimulus-response" versus "cognitive" memory in animals. He argues that the correspondence metaphor is well suited to capture explicit/"cognitive" memory phenomena in both domains, but may be less useful for implicit/stimulus-response type phenomena.

In contrast to these commentators, however, **Neisser** and **Winograd**, while also emphasizing the functional-behavioral aspects of memory, seem to ground their functional perspective in a humanistic-social view of man, focussing on more explicit and controlled forms of remembering. For **Neisser**, "remembering is a land of doing. Like other kinds of doing, it is purposive, personal, and particular." This view is also shared by **Winograd**, who stresses the social function of remembering. Both **Neisser** and **Winograd** emphasize that memory-based behaviors, such as joke or story telling, or even sharing personal memories, may have other goals apart from accurate reproduction, such as entertaining or impressing others. The personal and social goals that they emphasize (e.g., impression management) are not unique to "remembering," but are subserved by other behaviors as well (compare **Kruglanski's** concern that the correspondence metaphor may leave out aspects of cognitive activity that are unique to "memory"). Likewise, according to **Anderson**, "the function of memory is to make past experience useful in pursuit of present goals." More specifically, "the typical goal involves value judgments: approach-avoid, good-bad, etc." His view also emphasizes operations, valuation, and integration, which are "fundamental to function in general and to memory in specific."

We are pleased that our article has stimulated a critical discussion of the function of memory in everyday life, and are sympathetic to the concerns brought out in these comments. Like these other proposals, the correspondence metaphor is also motivated by an important function of memory in everyday life: that of providing a faithful representation of past events. Thus, we emphasized that in this conception, "memory does not serve merely as a depository of isolated, lifeless units, but rather affords a meaningful representation of real-life events that can be effectively utilized in future interactions" (sect. 2.2). We also stressed that for this very reason, the evaluation of memory correspondence must also take functional considerations into account in weighting the different aspects or dimensions of correspondence (e.g., accuracy vs. quantity, gist vs. detail; central vs. peripheral information; see sects. 4.1 and 5.3.2).

Of course, we agree that memory in real life serves functions other than that of providing accurate information about the past. In fact, it is rarely the case that accurate reproduction (or correspondence) is a goal in itself. Even in the most artificial of laboratory experiments (cf. **Winograd**), accurate remembering is probably subordinate to other goals, such as maintaining one's self-esteem, impressing the experimenter, and so forth. How much more so in natural situations! Nevertheless, even if we consider such everyday goals as impression management (**Neisser**, **Winograd**) or attitude formation (**Anderson**), we think it is clear that such goals will also generally be served by having available an accurate representation of the past, whether or

not that representation is put to use. (Consider also the importance that people attach to the validation of their own and other peoples memories; **Begg, Newby & Ross.**) Thus, unlike Neisser and Winograd, who cast correspondence (verity) and utility as two ends of a continuum (emphasizing situations in which they diverge), we see correspondence as generally, though not always, subvenient to utility.

In sum, memory is not monolithic, and any attempt to characterize it in terms of a single quality or function will certainly not do justice to its inherent heterogeneity (cf. **Anderson's** claim that "this [valuation and integration] is what memory is for"). Indeed, any single metaphor, correspondence included, is likely to capture but a limited part of memory's full nature. Hence, in order to encompass the many facets of memory, we hope that much more versatility will be seen in the use of memory metaphors and in developing their ensuing research orientations than has been witnessed in the first hundred years or so of memory research.

## References

**Letter "a" and \*V appearing before authors\* initials refer to target article and response respectively.**

- Aanstoos, C. M. (1991) Experimental psychology and the challenge of real life. *American Psychologist* 46:77-78. [aAK]
- Alba, J. W. & Hasher, L. (1983) Is memory schematic? *Psychological Bulletin* 93:203-31. [aAK]
- Algom, D. (1992) Memory psychophysics: An examination of its perceptual and cognitive prospects. In: *Psychophysical approaches to cognition*, ed. D. Algom. Elsevier. [aAK]
- Algom, D. & Cain, W. S. (1991) Remembered odors and mental mixtures: Tapping reservoirs of olfactory knowledge. *Journal of Experimental Psychology: Human Perception and Performance* 17:1104-19. [aAK]
- Algom, D. & Lubel, S. (1994) Psychophysics in the field: Perception and memory for labor pain. *Perception* 23:133-141. [DA]
- Algom, D., Wolf, Y. & Bergman, B. (1985) Integration of stimulus dimensions in perception and memory: Composition rules and psychophysical relations. *Journal of Experimental Psychology: General* 114:451-71. [aAK]
- Allen, G. L., Siegel, A. W. & Rosinski, R. R. (1978) The role of perceptual context in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning and Memory* 4:617-30. [aAK]
- Allport, G. W. (1961) *Patterns and growth in personality*. Holt, Rinehart and Winston. [aAK]
- Alterman, R. (1988) Adaptive planning. *Cognitive Science* 12:393-421. [RA]
- Alterman, R., Zito-Wolf, R. & Carpenter, T. (1991) Interaction, comprehension, and instruction usage. *Journal of the Learning Sciences* 1(4). [RA]
- (1995) *Pragmatic action* [Technical Report CS-95-180]. Brandeis University. [RA]
- Anderson, J. R. (1983) *The architecture of cognition*. Harvard University Press. [TPM]
- (1991) Reflections of the environment in memory. *Psychological Science* 2:396-408. [UN]
- Anderson, N. H. (1981) *Foundations of information integration theory*. Academic Press. [NHA]
- (1991) Functional memory in person cognition. In: *Contributions to information integration theory: vol. 1. Cognition*, ed. N. H. Anderson. Erlbaum. [NHA]
- (1992) Integration psychophysics and cognition. In: *Psychophysical approaches to cognition*, ed. D. Algom. Elsevier. [NHA]
- (in press) *A functional theory of cognition*. Erlbaum. [NHA]
- Arbib, M. A. & Hesse, M. B. (1986) *The construction of reality*. Cambridge University Press. [aAK]
- Arnold, M. B. (1984) *Memory and the brain*. Erlbaum. [AB-Z]
- Atkinson, R. C. & Shiffrin, R. M. (1968) Human memory: A proposed system and its control processes. *The psychology of learning and motivation: advances in research and theory*, vol. 2, ed. K. W. Spence. Academic Press. [aAK]
- Bacon, F. T. (1979) Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory* 5:241-52. [1MB]
- Baddeley, A. D. (1988) But what the hell is it for? In *Practical aspects of memory: Current research and issues*, vol. 1, ed. M. Gruneberg, P. Morris, & R. Sykes. Wiley. [aAK, EW]
- (1990) *Human memory: Theory and practice*. Erlbaum. [aAK]
- Baddeley, A. D., Lewis, V. & Nimmo-Smith, I. (1978) When did you last . . . ? In: *Practical aspects of memory*, ed. M. Gruneberg, P. Morris & R. Sykes. Academic Press. [aAK]
- Bahrick, H. P. (1984a) Associations and organization in cognitive psychology: A reply to Neisser. *Journal of Experimental Psychology: General* 113:36-37. [HPB]
- (1984b) Memory for people. In: *Everyday memory, actions and absentmindedness*, ed. J. E. Harris & P. E. Morris. Academic Press. [EW]
- (1987) Functional and cognitive memory theory: An overview of some key issues. In: *Memory and learning: The Ebbinghaus centennial conference*, ed. D. S. Gorfein & R. R. Hoffman. Erlbaum. [aAK]
- (1992) Stabilized memory of unrehearsed knowledge. *Journal of Experimental Psychology: General* 121:112-13. [aAK]
- Bahrick, H. P. & Bahrick, P. O. (1994) A re-examination of the interrelations among measures of retention. *Quarterly Journal of Experimental Psychology* 16:318-24. [LK]
- Bahrick, H. P., Bahrick, P. O. & Wittlinger, R. P. (1975) Fifty years of memories for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General* 104:54-75. [EW]
- Bahrick, H. P., Hall, L. K. & Berger, S. A. (in press) Accuracy and distortion in memory for high school grades. *Psychological Science*. [HPB]
- Bahrick, H. P., Hall, L. K. & Dunlosky, J. (1993) Reconstructive processing of memory content for high versus low test scores and grades. *Applied Cognitive Psychology* 7:1-10. [aAK]
- Baker-Ward, L., Ornstein, P. A. & Gordon, B. N. (1993) A tale of two settings: Young children's memory performance in the laboratory and the field. In: *Memory in everyday life*, ed. G. M. Davies & R. H. Logie. Elsevier. [aAK]
- Ballard, D. H., Hayhoe, M. M., Li, F. & Whitehead, S. D. (1992) Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London* 337:331-39. [KK]
- Ballard, D. H., Hayhoe, M. M. & Pelz, J. (1994a) Memory limits in sensory-motor tasks. In: *Models of information processing in the basal ganglia*, ed. J. C. Houk, J. L. Davis & D. G. Beiser. MIT Press. [KK]
- (1994b) Memory representations in natural tasks. *Journal of Cognitive Neuroscience* 7(1):66-80. [KK]
- Banaji, M. R. & Crowder, R. (1989) The bankruptcy of everyday memory. *American Psychologist* 44:1185-93. [aAK, LK, UN, DBW]
- Banks, W. P. (1970) Signal detection theory and human memory. *Psychological Bulletin* 74:81-99. [aAK]
- Barclay, C. R. (1986) Schematization in autobiographical memory. In: *Autobiographical memory*, ed. D. C. Rubin. Cambridge University Press. [aAK]
- (1988) Truth and accuracy in autobiographical memory. In: *Practical aspects of memory: Current research and issues*, vol. 2, ed. M. Gruneberg, P. Morris & R. Sykes. Wiley. [aAK, JPS]
- (1993) Remembering ourselves. In: *Memory in everyday life*, ed. G. M. Davies & R. H. Logie. Elsevier. [aAK]
- Barclay, C. R. & Wellman, H. M. (1986) Accuracies and inaccuracies in autobiographical memories. *Journal of Memory and Language* 25:93-103. [aAK, JPS]
- Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzone, G. & Narens, L. (1995) A model of metamemory components involved in retrieval, (submitted). [aAK]
- Bartlett, F. C. (1932) *Remembering*. Cambridge University Press. [aAK, NHA, HPB]
- Bechtel, W. (1988) *Philosophy of mind: An overview for cognitive science*. Erlbaum. [aAK]
- Bechtel, W. & Abrahamsen (1991) *Connectionism and the mind: An introduction to parallel processing in networks*. Blackwell. [aAK]
- Begg, I. M., Anas, A. & Faranacci, S. (1992) Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General* 121:446-58. [1MB]
- Begg, I. M. & Armour, V. (1991) Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioral Science* 22:195-313. [1MB]
- Begg, I. M., Armour, V. & Kerr, T. (1985) On believing what we remember. *Canadian Journal of Behavioral Science* 23:199-214. [1MB]
- Begg, I. M., Duft, S., Lalonde, P., Melnick, R. & Sanvito, J. (1989) Memory predictions are based on ease of processing. *Journal of Memory and Language* 28:610-32. [aAK]
- Begg, I. M., Robertson, R. K., Gruppuso, V., Anas, A. & Needham, D. R. (in

- press) The illusory knowledge effect. *Journal of Memory and Language*. [IMB]
- Bekerian, D. A. & Bowers, J. M. (1983) Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Human Learning and Memory* 9:139-45. [MK]
- Ben-Ze'ev, A. (1986) Two approaches to memory. *Philosophical Investigations* 9:288-301. [AB-Z] (1993) *The perceptual system: A philosophical and psychological perspective*. Lang. [AB-Z]
- Bensinger, D. G., Hayhoe, M. & Ballard, D. (1995) Visual memory in a natural task. *Investigative Ophthalmology and Visual Science* [Suppl.] 36(4):S14. [KK]
- Bernbach, H. A. (1967) Decision processes in memory. *Psychological Review* 74:462-80. [aAK]
- Biederman, I. (1987) Recognition by components: A theory of human image understanding. *Psychological Review* 94:115-45. [GM]
- Bjork, R. A. (1994) Memory and metamemory considerations in the training of human beings. In: *Metacognition: Knowing about knowing*, ed. J. Metcalfe & A. P. Shimamura. MIT Press. [aAK]
- Black, M. (1962) *Models and metaphors*. Cornell University Press. [aAK]
- Boon, J. C. W. & Davies, G. (1988) Attitudinal influences on witness memory: Fact and fiction. In: *Practical aspects of memory: Current research and issues*, vol. 1, ed. M. Gruneberg, P. Morris & R. Sykes. Wiley. [aAK]
- Bousfield, W. A. (1953) The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology* 49:229-40. [aAK]
- Bousfield, W. A. & Rosner, S. R. (1970) Free vs. uninhibited recall. *Psychonomic Science* 20:75-76. [aAK]
- Bower, G. H. (1970) Organizational factors in memory. *Cognitive Psychology* 1:18-46. [aAK]
- Bower, G. H., Black, J. & Turner, T. (1979) Scripts in text comprehension and memory. *Cognitive Psychology* 11:177-220. [aAK]
- Bransford, J. D. & Franks, J. J. (1971) The abstraction of linguistic ideas. *Cognitive Psychology* 2:331-50. [aAK]
- Bransford, J. D. & Johnson, M. K. (1972) Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior* 11:717-26. [aAK]
- Bransford, J. D., McCarrell, N. S., Franks, J. J. & Nitsch, K. E. (1977) Toward unexplaining memory. In *Perceiving, acting, and knowing*, ed. R. Shaw & J. D. Bransford. Erlbaum. [aAK]
- Breuer, J. & Freud, S. (1895/1955) Studies on hysteria. In: *The standard edition of the complete psychology works of Sigmund Freud* (2 vols.), ed. J. Strachey. Hogarth Press. [MAC]
- Brewer, W. F. (1988) Memory for randomly sampled autobiographical events. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK, EW] (1992) The theoretical and empirical status of the flashbulb memory hypothesis. In: *Affect and accuracy in recall: Studies of "flashbulb memories"*, ed. E. Wingrad & U. Neisser. Cambridge University Press. [UR] (in press) What is recollective memory? In: *Remembering our past: Studies in autobiographical memory*, ed. D. C. Rubin. Cambridge University Press. [aAK]
- Brewer, W. F. & Nakamura, G. U. (1984) The nature and function of schemas. In: *Handbook of social cognition*, ed. R. S. Wyer & T. K. Srull. Erlbaum. [aAK]
- Brewer, W. F. & Treyners, J. C. (1981) Role of schemata in memory for places. *Cognitive Psychology* 13:207-30. [aAK]
- Broadbent, D. E. (1971) *Decision and stress*. Academic press. [aAK]
- Brown, A. L. (1979) Theories of memory and the problem of development: Activity, growth, and knowledge. In: *Levels of processing*, ed. F. I. M. Craik & L. Cermak. Erlbaum. [rAK]
- Brown, E. L., Deffenbacher, K. A. & Sturgill, W. (1977) Memory for faces and the circumstances of encounter. *Journal of Applied Psychology* 62:311-15. [aAK]
- Brown, J. (Ed.) (1976) *Recall and recognition*. Wiley. [aAK]
- Brown, R. & Kulik, J. (1977) Flashbulb memories. *Cognition* 5:73-99. [DBW]
- Bruce, D. (1985) The how and why of ecological memory. *Journal of Experimental Psychology: General* 114:78-90. [aAK] (1989). Functional explanations of memory. In: *Everyday cognition in adulthood and late life*, ed. L. W. Poon, D. C. Rubin & B. E. Wilson. Cambridge University Press. [aAK] (1991) Mechanistic and functional explanations of memory. *American Psychologist* 46:46-49. [aAK]
- Budescu, D. & Bar-Hillel, M. (1993) To guess or not to guess: A decision-theoretic view of formula scoring. *Journal of Educational Measurement* 38:277-291. [aAK]
- Bunsey, M. & Eichenbaum, H. (in press a) Selective damage to the hippocampal region blocks long-term retention of a natural and nonspatial stimulus-stimulus association. *Hippocampus*. [HE] (in press b) Conservation of hippocampal memory function in rats and humans. *Nature*. [HE]
- Bursen, H. A. (1978) *Dismantling the memory machine*. Reidel. [AB-Z]
- Byrne, R. W. (1979) Memory for urban geography. *Quarterly Journal of Experimental Psychology* 31:147-54. [aAK]
- Carruthers, M. (1990) *The book of memory: A study of memory in medieval culture*. Cambridge University Press. [TON]
- Ceci, S. J. & Bronfenbrenner, U. (1985) Don't forget to take the cup cakes out of the oven: Prospective memory, strategic time monitoring, and context. *Child Development* 56:152-64. [aAK] (1991) On the demise of everyday memory: "The rumors of my death are much exaggerated" (Mark Twain). *American Psychologist* 46:27-32. [aAK]
- Ceci, S. J. & Bruck, M. (1993) The suggestibility of the child witness: An historical review and synthesis. *Psychological Bulletin* 113:403-39. [rAK]
- Chambers, D. & Reisberg, D. (1992) What an image depicts depends on what an image means. *Cognitive Psychology* 24:165-76. [GM]
- Christie, D. F. M. & Ellis, H. D. (1980) Photofit constructions versus verbal descriptions of faces. *Journal of Applied Psychology* 66:358-63. [aAK]
- Cohen, G. (1989) *Memory in the real world*. Erlbaum. [aAK]
- Cohen, N. J. & Eichenbaum, H. (1993) *Memory, amnesia, and the hippocampal system*. MIT Press. [HE]
- Cohen, R. L. (1988) Metamemory for words and enacted instructions: Predicting which items will be recalled. *Memory & Cognition* 16:452-60. [aAK]
- Conway, M. A. (1990) *Autobiographical memory: An introduction*. Buckingham: Open University Press. [MAC]
- (1991) In defense of everyday memory. *American Psychologist* 46:19-27. [aAK]
- (1992) A structural model of autobiographical memory. In: *Theoretical perspectives on autobiographical memory*, ed. M. A. Conway, D. C. Rubin, H. Spinnler & W. A. Wagenaar. Kluwer. [MAC]
- (1993) Method and meaning in memory research. In: *Memory in everyday life*, ed. G. M. Davies & R. H. Logie. Elsevier. [aAK, DBW]
- (1996) Autobiographical knowledge and autobiographical memories. In: *Remembering our past: Studies in autobiographical memory*, ed. D. C. Rubin. Cambridge University Press. [MAC]
- Conway, M. A., Cohen, G. & Stanhope, N. (1991) On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General* 120:395-409. [EW]
- Conway, M. A. & Rubin, D. C. (1993) The structure of autobiographical memory. In: *Theories of memory*, ed. A. E. Collins, S. E. Gathercole, M. A. Conway & P. E. M. Morris. Erlbaum. [MAC]
- Coren, S., Ward, L. M. & Enns, J. T. (1994) *Sensation and perception*. Harcourt Brace. [DA]
- Craik, F. I. M. (1983) On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London B*302:341-59. [aAK]
- (1983) On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London B*302:341-59. [AWK]
- (1991) Will cognitivism bury experimental psychology? *Canadian Psychology* 32:440-44. [aAK]
- Craik, F. I. M. & Lockhart, R. S. (1972) Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11:671-84. [aAK]
- Cronbach, L. J. (1955) Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin* 52:177-93. [aAK]
- (1957) The two disciplines of scientific psychology. *American Psychologist* 12:671-84. [TPM, DBW]
- (1984) *Essentials of psychological testing*. Harper & Row. [aAK]
- Crowder, R. G. (1976) *Principles of learning and memory*. Erlbaum. [aAK]
- (1993) Systems and principles in memory theory: Another critique of pure memory. In: *Theories of memory*, ed. A. F. Collins, S. E. Gathercole, M. A. Conway & P. E. M. Morris. Erlbaum. [aAK]
- Davies, G. M. & Logie, R. H., Eds. (1993) *Memory in everyday life*. Elsevier. [aAK, LK]
- Dawes, R. M. (1966) Memory and distortion of meaningful written material. *British Journal of Psychology* 57:77-86. [aAK]
- Deffenbacher, K. A. (1988) Eyewitness research: The next ten years. In: *Practical aspects of memory: Current research and issues*, vol. 1, ed. M. Gruneberg, P. Morris & R. Sykes. Wiley. [aAK] (1991) A maturing of research on the behavior of eyewitnesses. *Applied Cognitive Psychology* 5:377-402. [aAK]
- Donahoe, J. W. & Palmer, D. C. (1994) *Learning and complex behavior*. Allyn & Bacon. [DCP]

## References/Koriat & Goldsmith: Memory metaphors

- Dooling, D. J. & Christiaansen, R. E. (1977) Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory* 3:428-36. [aAK, HPB]
- Dooling, D. J. & Mullet, R. L. (1973) Locus of thematic effects in retention of prose. *Journal of Experimental Psychology* 97:404-6. [aAK]
- Duggan III, L. M., Aubrey, M., Doherty, E., Isquith, P., Levine, M. & Scheiner, J. (1989) The credibility of children as witnesses in a simulated child sexual abuse trial. In: *Perspectives on children's testimony*, ed. S. J. Ceci, D. F. Ross & M. P. Toglia. Springer-Verlag. [rAK]
- Ebbinghaus, H. (1895/1964) *Memory: A contribution to experimental psychology*. Dover. [aAK, MAC]
- Edwards, D. & Potter, J. (1992) The chancellor's memory: Rhetoric and truth in discursive remembering. *Applied Cognitive Psychology* 6:187-215. [UN]
- Eich, J. E. (1980) The cue-dependent nature of state-dependent retrieval. *Memory & Cognition* 8:157-73. [aAK]
- Eichenbaum, H., Otto, T. & Cohen, N. J. (1994) Two functional components of the hippocampal memory system. *Behavioral and Brain Sciences* 17:449-518. [HE]
- Ellis, H. D., Shepard, J. W. & Davies, G. M. (1975) An investigation of the use of the photofit technique for recalling faces. *British Journal of Psychology* 66:29-37. [aAK]
- Erdelyi, M. H. & Becker (1974) Hypernesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology* 6:159-71. [aAK]
- Erdelyi, M. H., Finks, J. & Feigin-Pfau, M. B. (1989) The effect of response bias on recall performance, with some observations on processing bias. *Journal of Experimental Psychology: General* 118:245-54. [aAK]
- Fechner, G. T. (1882/1987) Some thoughts on the psychophysical representation of memories (E. Scheerer, Trans.). *Psychological Research* 49:209-12. [aAK]
- Fischhoff, B., Slovic, P. & Lichtenstein, S. (1977) Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance* 3:552-64. [aAK]
- Fisher, R. P. (1981) The interaction between encoding distinctiveness and test conditions. *Journal of Experimental Psychology: Human Learning and Memory* 7:306-10. [aAK] (in press) Interviewing victims and witnesses of crime. *Psychology, Public Policy, and Law*. [RPF]
- Fisher, R. P. & Craik, F. I. M. (1977) The interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning and Memory* 3:701-11. [aAK]
- Fisher, R. P. & Geiselman, R. E. (1992) *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Charles Thomas. [aAK]
- Fisher, R. P., Geiselman, R. E. & Amador, M. (1989) Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology* 74:722-27. [aAK, AWK] Fisher, R. P., Geiselman, R. E. & Raymond, D. S. (1987) Critical analysis of police interview techniques. *Journal of Police Science and Administration* 15:177-85. [aAK]
- Fivush, R. (1988) The functions of event memory: Some comments on Nelson and Barsalou. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK] (1993) Commentary: Memory and meaning in the laboratory and in the field. In: *Memory in everyday life*, ed. G. M. Davies & R. H. Logie. Elsevier. [aAK]
- Fodor, J. A. (1975) *The language of thought*. Harvester Press. [aAK]
- Freud, S. (1898) Zum psychischen Mechanismus der Vergesslichkeit. *Monatsschrift für Psychiatrie und Neurologie* 1:436-43. [HJM] (1900/1953) The interpretation of dreams. In: *The standard edition of the complete psychological works of Sigmund Freud* (vols. 4 & 5), ed. J. Strachey. Hogarth Press. [MAC] (1905/1953) Fragment of an analysis of a case of hysteria. In: *The standard edition of the complete works of Sigmund Freud*, vol. 7, ed. J. Strachey. Hogarth Press. [SFL]
- Funder, D. C. (1987) Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin* 101:75-90. [aAK]
- Gardiner, J. M. & Java, R. I. (1993) Recognizing and remembering. In: *Theories of memory*, ed. A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris. Erlbaum. [GM]
- Garner, W. R. (1972) The acquisition and application of knowledge: A symbiotic relation. *American Psychologist* 27:941-46. [TPM]
- Geiselman, R. E. & Fisher, R. P. (1989) The cognitive interview technique for interviewing victims and witnesses of crime. In: *Psychological methods in investigation and evidence*, ed. D. Raskin. Springer. [aAK]
- Gentner, D. & Grudin, J. (1985) The evolution of mental metaphors in psychology: A 90-year retrospective. *American Psychologist* 40:181-92. [aAK]
- Gergen, K. J. (1994) Mind, text, and society: Self-memory in social context. In: *The remembering self Construction and accuracy in the self-narrative*, ed. U. Neisser & R. Fivush. Cambridge University Press. [UN]
- Gibson, J. J. (1979) *The ecological approach to visual perception*. Houghton Mifflin. [aAK, KGW, AB-Z]
- Gigerenzer, G., Hoffrage, U. & Kleinbolting, H. (1991) Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review* 98:506-28. [aAK]
- Gillund, G. & Shiffrin, R. M. (1984) A retrieval model for both recognition and recall. *Psychological Review* 91:1-67. [TPM]
- Goldmeier, E. (1982) *The memory trace: Its formation and its fate*. Erlbaum. [aAK]
- Goldstein, E. B. (1989) *Sensation and perception*. Wadsworth. [DA]
- Gorenstein, G. W. & Ellsworth, P. C. (1980) Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology* 65:616-22. [aAK]
- Gould, J. L. & Marler, P. (1987) Learning by instinct. *Scientific American* 256:74-85. [HE]
- Green, P. (1993) Pericles put in his place.: The new CAH on the fifth century BC. *Times Literary Supplement* 4699 (April 23), p. 5. [JPS]
- Gregg, V. H. (1986) *Introduction to human memory*. Routledge & Kegan Paul. [aAK]
- Grossberg, S. (1987) Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science* 11:23-63. [aAK]
- Gruneberg, M. M. & Morris, P. E. (1992) Applying memory research. In: *Aspects of memory*, 2d ed., vol. 1, ed. M. Gruneberg & P. Morris. Routledge. [aAK]
- Gruneberg, M. M., Morris, P. E. & Sykes, R. N., Eds. (1988) *Practical aspects of memory: Current research and issues*. Wiley. [aAK]
- Gruneberg, M. M., Morris, P. E. & Sykes, R. N. (1991) The obituary on everyday memory and its practical applications is premature. *American Psychologist* 46:74-76. [aAK]
- Harnad, S. (1990) The symbol grounding problem. *Physica D* 42:335-46. [aAK] (1992) Connecting object to symbol in modeling cognition. In: *Connectionism in context*, ed. A. Clarke & R. Luce. Springer-Verlag. [aAK]
- Harris, J. E. & Morris, P. E., Eds. (1984) *Everyday memory, actions, and absent-mindedness*. Academic Press. [aAK]
- Hart, R. A. (1979) *Children's experience of place*. Irvington. [aAK] (1981) Children's spatial representation of the landscape: Lessons and questions from a field study. In: *Spatial representation and behavior across the life span*, ed. L. S. Liben, A. H. Patterson, & N. Newcombe. Academic Press. [aAK]
- Hastorf, A. H. & Cantril, H. (1954) They saw a game: A case study. *Journal of Abnormal and Social Psychology* 49:129-134. [IRN]
- Hayhoe, M. M., Ballard, D. H. & Pelz, J. B. (1994) Visual representations in natural tasks. In: *Proceedings of the workshop on visual behaviors*. Los Alamitos, CA: IEEE Computer Society Press. [KK]
- Head, H. (1920) *Studies in neurology*, vol. 2. Hodder & Stoughton and Oxford University Press. [aAK]
- Hermann, D. J. & Chaffin, R. (1988) *Memory in historical perspective*. Springer-Verlag. [aAK]
- Hesse, M. B. (1966) *Models and analogies in science*. University of Notre Dame Press. [aAK, TON]
- Hilgard, E. R. & Loftus, E. F. (1979) Effective interrogation of the eyewitness. *The International Journal of Clinical and Experimental Hypnosis* 27:342-57. [aAK]
- Hinton, G. E. & Anderson, J. A. (1981) *Parallel models of associative memory*. Erlbaum. [aAK]
- Hintzman, D. L. (1974) Psychology and the cow's belly. *The Worm Runner's Digest* 16:84-85. [aAK] (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 87:398-410. [aAK] (1990) Human learning and memory: Connections and dissociations. *Annual Review of Psychology*, 41:109-39. [aAK]
- Hoffman, R. R. (1980) Metaphor in science. In: *Cognitive and figurative language*, ed. R. P. Honeck & R. R. Hoffman. Erlbaum. [aAK]
- Holland, P. W. (1986) Statistics and causal inference. *Journal of the American Statistical Association* 81:945-60. [DBW]
- Huttenlocher, J., Hedges, L. V. & Bradburn, N. M. (1990) Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:196-213. [aAK]
- Huttenlocher, J., Hedges, L. V. & Duncan, S. (1991) Categories and particular! Prototype effects in estimating spatial location. *Psychological Review* 98:352-76. [aAK]

- Huttenlocher, J., Hedges, L. V. & Prohaska, V. (1988) Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review* 95:471-84. [aAK]
- Hyman, I. E., Jr. (1994) Conversational remembering: Story recall with a peer versus for an experimenter. *Applied Cognitive Psychology* 8:49-66. [UN]
- Hyman, I. E., Jr. & Neisser, U. (1992) The role of the self in recollections of a seminar. *Journal of Narrative and Life History* 2:81-103. [UN]
- Jacoby, L. L. (1988) Memory observed and memory unobserved. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK]
- Jacoby, L. L., Kelley, C. M. & Dywan, J. (1989) Memory attributions. In: *Varieties of memory and consciousness: Essays in honour of Endel Tulving*, ed. H. L. Roediger & F. I. M. Craik. Erlbaum. [aAK]
- Jacoby, L. L., Lindsay, D. S. 6c Toth, J. P. (1992) Unconscious influences revealed: Attention, awareness, and control. *American Psychologist* 47:802-9. [aAK]
- Janowsky, J. S., Shimamura, A. P. & Squire, L. R. (1989) Memory and metamemory: Comparisons between frontal lobe lesions and amnesic patients. *Psychobiology* 17:3-11. [aAK]
- Johnson, M. (1987) *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press. [UN]
- Johnson, M. K., Bransford, J. D. & Solomon, S. K. (1973) Memory for tacit implications of sentences. *Journal of Experimental Psychology* 98:203-5. [aAK]
- Johnson, M. K., Hastroudi, S. & Lindsay, D. S. (1993) Source monitoring. *Psychological Bulletin* 114:3-28. [MAC]
- Johnson-Laird, P. N. (1983) *Mental models*. Harvard University Press. [aAK]
- Jussim, L. (1991) Social perception and social reality: A reflection construction model. *Psychological Review* 98:54-73. [aAK]
- Karn, K.S. (1995) *Spatial representations for programming saccadic eye movements*. Doctoral dissertation, University of Rochester, Rochester, NY. [KK]
- Kelley, C. M. & Jacoby, L. L. (1990) The construction of subjective experience: Memory attributions. *Mind and Language* 5:49-68. [aAK]
- Kelley, C. M. & Lindsay, D. S. (1993) Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language* 32:1-24. [aAK]
- Kelly, G. A. (1955) *The psychology of personal constructs*, vol. 1 & 2. Norton. [aAK]
- Kenny, D. A. & Albright, L. (1987) Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin* 102:390-402. [aAK]
- Kenny, D. A. (1991) A general model of consensus and accuracy in interpersonal perception. *Psychological Review* 98:155-63. [aAK, AWK]
- Kerst, S. M. & Howard, J. H. (1978) Memory psychophysics for visual area and length. *Memory & Cognition* 6:327-35. [aAK]
- Kihlstrom, J. F. (1994) Memory research: The convergence of theory and practice. Paper presented at the Third Practical Aspects of Memory Conference (August 1994) College Park, MD. [LK]
- Kintsch, W. (1967) Memory and decision aspects of recognition learning. *Psychological Review* 74:496-504. [aAK]
- Kintsch, W. & van Dijk, T. A. (1978) Toward a model of text comprehension and production. *Psychological Review* 85:363-94. [aAK]
- Klatzky, R. L. (1984) *Memory and awareness: An information-processing perspective*. Freeman. [RPF]
- (1991) Lets be friends. *American Psychologist* 46:43-46. [aAK]
- Klatzky, R. L. & Erdelyi, M. H. (1985) The response criterion problem in tests of hypnosis and memory. *International Journal of Clinical and Experimental Hypnosis* 33:246-57. [aAK]
- Kolers, P. A. (1973) Remembering operations. *Memory & Cognition* 1:347-55. [aAK]
- Kolers, P. A. & Roediger, H. L. (1984) Procedures of mind. *Journal of Verbal Learning and Verbal Behavior* 23:425-49. [aAK]
- Kolers, P. A. & Smythe, W. E. (1984) Symbol manipulation: Alternatives to the computational view of mind. *Journal of Verbal Learning and Verbal Behavior* 23:289-314. [aAK]
- Koriat, A. (1993) How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review* 100:609-39. [aAK]
- (1995) Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General* 124:311-33. [aAK, BLS]
- Koriat, A., Fischhoff, B. & Razel, O. (1976) An inquiry into the process of temporal orientation. *Acta Psychologica* 40:57-73. [aAK]
- Koriat, A. & Goldsmith, M. (1994) Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General* 123:297-316. [aAK, RPF, AWK]
- (in press a) Memory as something that can be counted versus memory as something that can be counted on. In: *Basic and applied memory research-Practical applications*, vol. 2, ed. D. Hermann, C. McEvoy, C. Herzog, S. M. Johnson & P. Hertel. Erlbaum. [aAK] (in press b) Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*. [aAK]
- Kozminsky, E. (1977) Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition* 5:482-90. [aAK]
- Kruglanski, A. W. (1989) The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin* 106:395-409. [aAK, AWK]
- Kuhn, T. S. (1979) *Metaphor in science*. In: *Metaphor and thought*. Cambridge University Press. [aAK]
- Kvavilashvili, L. & Ellis, J. A. (1995) Ecological validity and the convergence of traditional and ecological approaches to memory research (submitted). [LK]
- Lachman, R. (1960) The model in theory construction. *Psychological Review* 67:113-29. [TON]
- Lachman, R., Lachman, J. & Butterfield, E. C. (1979) *Cognitive psychology and information processing: An introduction*. [aAK]
- Lakoff, G. (1987) *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press. [UR]
- Laming, D. (1992) Analysis of short-term retention: Models for Brown-Peterson experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:1342-1365. [KGW]
- Landauer, T. K. (1975) Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology* 7:495-531. [aAK]
- Larsen, S. F. (1987) Remembering and the archaeology metaphor. *Metaphor and Symbolic Activity* 2:187-99. [SFL] (1988) Remembering reported events: Memory for news in ecological perspective. In: *Practical aspects of memory: Current research and issues*, vol. 1, ed. M. Gruneberg, P. Morris, & R. Sykes. Wiley. [aAK]
- Lave, J. (1988) *Cognition in practice*. Cambridge University Press. [RA]
- Leary, D. E., Ed. (1990) *Metaphors in the history of psychology*. Cambridge University Press. [aAK]
- Lichtenstein, S., Fischhoff, B. & Phillips, L. D. (1982) Calibration of probabilities: The state of the art to 1980. In: *Judgment under uncertainty: Heuristics and biases*, D. Kahneman, P. Slovic & A. Tversky. Cambridge University Press. [aAK]
- Lindsay, R. C. L. & Wells, G. L. (1985) Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology* 70:556-64. [aAK]
- Linton, M. (1975) Memory for real-world events. In: *Explorations in cognition*, ed. D. A. Norman & D. Rumelhart. Freeman. [aAK]
- Lipton, J. P. (1977) On the psychology of eyewitness testimony. *Journal of Applied Psychology* 62:90-95. [aAK]
- Locke, J. (1690/1965) *An essay concerning human understanding*. Collier Books. [aAK]
- Lockhart, R. S. & Craik, F. I. M. (1990) Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology* 44:87-112. [aAK]
- Lockhart, R. S. & Murdock, B. B. (1970) Memory and the theory of signal detection. *Psychological Bulletin* 74:100-109. [aAK]
- Loftus, E. F. (1977) How to catch a zebra in semantic memory. In: *Perceiving, acting and knowing*, ed. R. Shaw & J. Bransford. Erlbaum. [aAK]
- (1979a) *Eyewitness testimony*. Harvard University Press. [aAK]
- (1979b) The malleability of memory. *American Scientist* 67:312-70. [aAK] (1982) Memory and its distortions. In: *G. Stanley Hall Lectures*, ed. A. G. Kraut. American Psychological Association. [aAK] (1991a) Made in memory: Distortions to recollection after misleading information. In: *Psychology of learning and motivation*, vol. 27, ed. G. Bower. Academic Press. [GM] (1991b) The glitter of everyday memory . . . and the gold. *American Psychologist* 46:16-19. [aAK]
- Loftus, E. F. & Hoffman, H. G. (1989) Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General* 118:100-104. [aAK]
- Loftus, E. F. & Marburger, W. (1983) Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition* 11:114-20. [aAK]
- Loftus, E. F. & Palmer, J. C. (1974) Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior* 13:585-89. [aAK]
- Loftus, E. F., Miller, D. G. & Burns, H. J. (1978) Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory* 4:19-31. [aAK]

## References/Koriat & Goldsmith: Memory metaphors

- Loftus, G. R. (1985) Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11:397-406. [KGW]
- Lord, A. B. (1960) *The singer of tales*. Harvard University Press. [UN]
- Malcolm, N. (1977) *Memory and mind*. Cornell University Press. [aAK]
- Mandler, J. M. (1979) Categorical and schematic organization in memory. In: *Memory organization and structure*, ed. C. R. Puff. Academic Press. [aAK]
- (1992) How to build a baby: 2. Conceptual primitives. *Psychological Review* 99:587-604. [UN]
- Mandler, J. M. & Johnson, N. S. (1977) Remembrance of things parsed: Story structure and recall. *Cognitive Psychology* 9:111-51. [aAK]
- Mandler, J. M. & Mandler, G. (1964) *Thinking: From association to gestalt*. Wiley. [aAK]
- Mantyla, T. (1986) Optimizing cue effectiveness: Recall of 500 and 600 incidentally learned words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12:66-71. [aAK]
- Markus, H. (1977) Self-schemata and processing information about the self. *Journal of Personality and Social Psychology* 35:63-78. [aAK] (1980) The self in thought and memory. In: *The self in social psychology*, ed. D. M. Wegner & R. R. Vallacher. Oxford University Press. [aAK]
- Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. Freeman. [GM]
- Marshall, J. C. & Fryer, D. M. (1978) Speak, memory! An introduction to some historic studies of remembering and forgetting. In: *Aspects of memory*, ed. M. M. Gruneberg & P. Morris. Methuen. [aAK]
- McBurney, D. H. & Collings, V. B. (1977) *Introduction to sensation/perception*. Prentice-Hall. [DA]
- McCauley, R. N. (1988) Walking in our own footsteps: Autobiographical memory and reconstruction. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK]
- McClelland, J. L. & Rumelhart, D. E. (1985) Distributed memory and the representation of general and specific knowledge. *Journal of Experimental Psychology: General* 114:159-88. [aAK] (1986) A distributed model of human learning and memory. In: *Parallel distributed processing: Explorations in the microstructure of cognition: vol. 2. Psychological and biological models*, ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [aAK]
- McDonald, R. J. & White, N. M. (1995) Hippocampal and nonhippocampal contributions to place learning in rats. *Behavioral Neuroscience* 109:579-93. [HE]
- McNamara, T. P. (1986) Mental representations of spatial relations. *Cognitive Psychology* 18:87-121. [aAK]
- Metcalfe, J. (1985) levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review* 92:1-38. [aAK]
- Metcalfe, J. (1990) Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General* 119:145-60. [aAK]
- (1993) Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review* 100:3-22. [aAK]
- Metcalfe, J., Schwartz, B. L. & Joaquim, S. G. (1993) The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:851-61. [aAK]
- Metcalfe, J. & Shimamura, A. (1994) *Metacognition: Knowing about knowing*. MIT Press. [aAK, TON]
- Minsky, M. (1975) A framework for representing knowledge. In: *The psychology of computer vision*, ed. P. H. Winston. McGraw-Hill. [aAK]
- Morris, C. D., Bransford, J. D. & Franks, J. J. (1977) Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior* 16:519-33. [aAK, EW]
- Morris, C. D., Stein, B. S. & Bransford, J. D. (1979) Prerequisites for the utilization of knowledge in the recall of prose passages. *Journal of Experimental Psychology: Human Learning and Memory* 5:253-61. [aAK]
- Morris, P. E., Tweedy, M. & Gruneberg, M. M. (1985) Interest, knowledge and the memory of soccer scores. *British Journal of Psychology* 76:417-25. [aAK]
- Morton, J. (1991) The bankruptcy of everyday thinking. *American Psychologist* 46:32-34. [aAK]
- Moscovitch, M. (1994) Memory and working with memory: Evaluation of a component process model and comparisons with other models. In *Memory Systems 1994*, ed. D. Schacter & E. Tulving. MIT Press. [BLS] (in press) Confabulation. In: *Memory distortion*, ed. D. L. Schacter, J. T. Coyle, G. D. Fischbach, M. M. Mesulam & L. E. Sullivan. Harvard University Press. [aAK]
- Moston, S. (1987) The suggestibility of children in interview studies. *First Language* 7:67-78. [aAK]
- Moyer, R. S. (1973) Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics* 13:180-84. [aAK]
- Moyer, R. S. & Dumais, S. T. (1978) Mental comparison. In: *The psychology of learning and motivation*, ed. G. H. Bower. Academic Press. [aAK]
- Murdock, B. B. (1966) The criterion problem in short-term memory. *Journal of Experimental Psychology* 72:317-24. [aAK] (1974) *Human memory: Theory and data*. Erlbaum. [aAK] (1982) A theory for the storage and retrieval of item and associative information. *Psychological Review* 89:609-26. [TPM]
- Murphy, A. H. (1973) A new vector partition of the probability score. *Journal of Applied Meteorology* 12:595-600. [aAK]
- Neisser, U. (1967) *Cognitive psychology*. Appleton-Century-Crofts. [aAK, SFL, UN]
- (1976) *Cognition and reality*. Freeman. [UN]
- (1978) Memory: What are the important questions? In: *Practical aspects of memory*, ed. M. M. Gruneberg, P. E. Morris & R. N. Sykes. Academic Press. [aAK, LK, AWK]
- (1981) John Dean's memory: A case study. *Cognition* 9:1-22. [aAK, JPS]
- (1984) Interpreting Harry Bahrick's discovery: What confers immunity against forgetting. *Journal of Experimental Psychology: General* 113:32-35. [aAK, HPB]
- (1986) Nested structure in autobiographical memory. In: *Autobiographical memory*, ed. D. C. Rubin. Cambridge University Press. [aAK]
- (1988a) The ecological approach to perception and memory. *New Trends in Experimental and Clinical Psychiatry* 4:153-66. [aAK] (1988b) Time present and time past. In: *Practical aspects of memory: Current research and issues*, vol. 2, ed. M. M. Gruneberg, P. Morris & R. Sykes. Wiley. [aAK, EW] (1988c) What is ordinary memory the memory of? In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK] (1988d) New vistas in the study of memory. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [LK] (1991) A case of misplaced nostalgia. *American Psychologist* 46:34-37. [aAK]
- Neisser, U. & Fivush, R., Eds. (in press) *The remembered self*. Cambridge University Press. [aAK]
- Neisser, U. & Harsch, N. (1992) Phantom flashbulbs: False recollections of hearing the news about Challenger. In: *Affect and accuracy in recall: Studies of "flashbulb memories"*, ed. E. Winograd & U. Neisser. Cambridge University Press. [aAK, UN]
- Neisser, U., Hyman, I. E., Jr., Harsch, N. & Usher, J. A. (in preparation) Accuracy and distortion in memories of a seminar. [UN]
- Neisser, U. & Winograd, E., Eds. (1988) *Remembering reconsidered: Ecological and traditional approaches to the study of memory*. Cambridge University Press. [aAK, UR]
- Neisser, U., Winograd, E., Bergman, E. T., Schreiber, C. A., Palmer, S. E. & Weldon, M. S. (in press) Remembering the earthquake: Direct experience vs. hearing the news. *Memory*. [UN]
- Neisser, U., Winograd, E. & Weldon, M. S. (1991) Remembering the earthquake: "What I experienced" versus "How I heard the news." Paper presented at the 32nd Annual Meeting of the Psychonomic Society, San Francisco, November, 1991. [aAK]
- Nelson, T. (1992) *Metacognition: Core readings*. Allyn & Bacon. [TON]
- (1996) *Consciousness and metacognition*. American Psychologist, in press. [TON]
- Nelson, T. O. & Chaiklin, S. (1980) Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning and Memory* 6:529-45. [aAK]
- Nelson, T. O. (1984) A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin* 95:109-33. [aAK]
- Nelson, T. O. & Narens, L. (1990) Metamemory: A theoretical framework and new findings. In: *The psychology of learning and motivation*, ed. G. Bower. American Psychologist. [aAK, GM, TON] (1994) Why investigate metacognition? In: *Metacognition: Knowing about knowing*, ed. J. Metcalfe & A. P. Shimamura. MIT Press. [aAK, TON]
- Neusner, J., Trans. (1984) Torah from our sages: Pirke Avot (reprinted and translated from the Hebrew). Chappaqua, NY: Rossel Books. [aAK]
- Newman, L. S. & Baumeister, R. F. (in press). Toward an explanation of the UFO abduction phenomenon: Hypnotic elaboration, extraterrestrial sadomasochism, and spurious memories. *Psychological Inquiry*. [IRN]
- Nigro, G. & Neisser, U. (1983) Point of view in personal memories. *Cognitive Psychology* 15:467-82. [aAK]
- Nilsson, L.-G. (1979) Functions of memory. In: *Perspectives on memory research: Essays in honor of Uppsala University's 500th anniversary*, ed. L.-G. Nilsson. Erlbaum. [aAK]
- (1987) Motivated memory: Dissociation between performance data and subjective reports. *Psychological Research* 49:183-88. [aAK, L-GN]
- Norman, D. A. & Wickelgren, W. A. (1969) Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology* 6:192-208. [aAK]

- O'Keefe, J. & Nadel, L. (1978) *The hippocampus as a cognitive map*. Oxford University Press. [HE]
- O'Neill, W. M. (1968) *The beginnings of modern psychology*. Penguin. [aAK]
- Oppenheimer, J. R. (1956) Analogy in science. *American Psychologist* 11:127-135. [aAK]
- Ornstein, P. A., Gordon, B. N. & Larus, D. M. (1992) Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology* 6:49-60. [aAK]
- Palmer, D. C. (1991) A behavioral interpretation of memory. In: *Dialogues on verbal behavior*, ed. L. J. Hayes & P. N. Chase. Context Press. [DCP]
- Pelz, J., Ballard, D. & Hayhoe, M. (1993) Memory use during performance of natural visuo-motor tasks. *Investigative Ophthalmology and Visual Science* [Suppl.] 34(4):1234. [KK]
- Peterson, L. R. & Peterson, M. J. (1959) Short-term retention of individual verbal items. *Journal of Experimental Psychology* 58:193-98. [KGW]
- Phillips, R. G. & LeDoux, J. E. (1991) Different contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral Neuroscience* 106:274-85. [HE]
- Pichert, J. W. & Anderson, R. C. (1977) Taking different perspectives on a story. *Journal of Educational Psychology* 69:309-15. [aAK]
- Pick, H. L. & Lockman, J. L. (1981) From frames of reference to spatial representations. In: *Spatial representation and behavior across the life span*, ed. L. S. Liben, A. H. Patterson & N. Newcombe. Academic Press. [aAK]
- Pipe, M-E. & Wilson, J. C. (1994) Cues and secrets: Influences on children's event reports. *Developmental Psychology* 30:515-25. [KGW]
- Pressley, M., Levin, J. R., Ghatala, E. S. & Ahmad, M. (1987) Test monitoring in young grade school children. *Journal of Experimental Child Psychology* 43:96-111. [aAK]
- Puff, C. R., Ed. (1982) *Handbook of research methods in human memory and cognition*. Academic Press. [aAK]
- R. v. *D.E.F.* [1995] A.J. No. 146 (Alberta Court of Appeal). [1MB]
- R. v. *Jmieff* [1994] B.C.J. No. 2506 (British Columbia Court of Appeal). [1MB]
- R. v. *S.C.H.* [1995] B.C.J. No. 237 (British Columbia Supreme Court). [1MB]
- Reder, L. M. & Ritter, F. E. (1992) What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:435-51. [aAK]
- Reid, T. (1785/1967) *Essays in the intellectual powers of man*. George Olms. [AB-Z]
- Riley, D. (1962) Memory for form. In: *Psychology in the making: Histories of selected research problems*, ed. L. Postman. Knopf. [aAK]
- Rock, I. (1983) *The logic of perception*. MIT Press. [aAK]
- Roediger, H. L. (1980) Memory metaphors in cognitive psychology. *Memory & Cognition* 8:231-46. [aAK, 1MB, SFL]
- (1991) They read an article? A commentary on the everyday memory controversy. *American Psychologist* 46:37-41. [aAK]
- Roediger, H. L. & McDermott, K. B. (1995) Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:803-14. [GM, EW]
- Roediger, H. L. & Payne, D. G. (1985) Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognition theories. *Memory & Cognition* 13:1-7. [aAK, GM]
- Roediger, H. L., Srinivas, K. & Waddil, P. (1989) How much does guessing influence recall? Comment on Erdelyi, Finks, and Feigin-Pfau. *Journal of Experimental Psychology: General* 118:253-57. [aAK]
- Roediger, H. L., Weldon, S. W. & Challis, B. H. (1989) Explaining dissociations between implicit and explicit measures of retention: A processing account. In: *Varieties of memory and consciousness: Essays in honour of Endel Tulving*, ed. H. L. Roediger & F. I. M. Craik. Erlbaum. [aAK]
- Ross, M. (1989) Relation of implicit theories to the construction of personal histories. *Psychological Review* 96:341-57. [aAK, AWK, UR, EW] (in press) Validating memories. In: *Memory for everyday and emotional events*, ed. N. L. Stein, P. A. Ornstein, B. Tversky & C. Brainerd. Erlbaum. [aAK, IRN]
- Ross, M. & Buehler, R. (1994) Creative remembering. In: *The remembering self: Construction and accuracy in the self-narrative*, ed. U. Neisser & R. Fivush. Cambridge University Press. [aAK]
- Ross, M. & Holmberg, D. (1990) Recounting the past: Gender differences in the recall of events in the history of a close relationship. In: *Self-inference processes: The Ontario Symposium*, vol. 6, ed. J. M. Olson & M. P. Zanna. Erlbaum. [IRN]
- (1992) Are wives' memories for events in relationships more vivid than their husbands' memories? *Journal of Social and Personal Relationships* 9:585-604. [IRN]
- Ross, M. & Newby, I. R. (in press). Distinguishing memory from fantasy. *Psychological Inquiry*. [IRN]
- Rubin, D. C. (1986) *Autobiographical memory*. Cambridge University Press. [aAK]
- (1995) *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press. [UN]
- Rumelhart, D. E. (1975) Notes on a schema for stories. In: *Representation and understanding*, ed. D. G. Bobrow & A. Collins. Academic Press. [aAK] (1980) Schemata: The building blocks of cognition. In: *Theoretical issues in reading comprehension*, ed. R. J. Spiro, B. C. Bruce, W. F. Brewer. Erlbaum. [aAK] (1989) The architecture of mind: A connectionist approach. In: *Foundations of cognitive science*, ed. M. I. Posner. MIT Press. [aAK]
- Rumelhart, D. E. & McClelland, J. L. (1986) *Parallel distributed processing*, vol. 1. MIT Press. [aAK]
- Rumelhart, D. E. & Norman, D. A. (1988) Representation in memory. In: *Handbook of experimental psychology*, ed. R. C. Atkinson, R. J. Herrnstein, G. Lindzey & R. D. Luce. Wiley. [aAK]
- Rumelhart, D. E., Smolensky, P., McClelland, J. L. & Hinton, G. E. (1986) Schemata and sequential thought processes in PDP models. In: *Parallel distributed processing*, vol. 2, ed. J. L. McClelland & D. E. Rumelhart. MIT Press. [aAK]
- Schacter, D. (1987) Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:501-18. [aAK] (1989) Memory. In: *Foundations of cognitive science*, ed. M. I. Posner. MIT Press. [aAK]
- Schank, R. C. (1982) *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press. [aAK]
- Schiffman, H. R. (1990) *Sensation and perception*. Wiley. [DA]
- Schlottmann, A. & Anderson, N. H. (1993) An information integration approach to phenomenal causality. *Memory & Cognition* 21:785-801. [NHA]
- Schneider, W. & Pressley, M. (1989) *Memory development between 2 and 20*. Springer-Verlag. [GM]
- Schwartz, B. L. (1994) Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review* 1:357-75. [aAK, BLS]
- Seifert, C. M., Robertson, S. P. & Black, J. B. (1985) Types of inferences generated during reading. *Journal of Memory and Language* 24:405-22. [aAK]
- Sekuler, R. & Blake, R. (1994) *Perception*. McGraw-Hill. [DA]
- Shepard, R. N. (1967) Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior* 6:156-63. [aAK] (1987) Toward a universal law of generalization for psychological science. *Science* 237:1317-23. [KGW]
- Sherry, D. F. & Schacter, D. L. (1987) The evolution of multiple memory systems. *Psychological Review* 94:439-54. [aAK]
- Shimamura, A. P. & Squire, L. R. (1986) Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12:452-60. [aAK, BLS]
- Siegel, A. W. (1981) The externalization of cognitive maps by children and adults: In search of ways to ask better questions. In: *Spatial representation and behavior across the life span*, ed. L. S. Liben, A. H. Patterson & N. Newcombe. Academic Press. [aAK]
- Siegel, A. W. & Schadler, M. (1977) Young people's cognitive maps of their classroom. *Child Development* 48:388-94. [aAK]
- Smith, W. M., Glenberg, A. & Bjork, R. A. (1978) Environmental context and human memory. *Memory & Cognition* 6:342-53. [aAK]
- Spence, D. P. (1982) *Narrative truth and historical truth*. Norton. [aAK]
- Spiro, R. J. (1980) Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior* 19:84-95. [aAK] *State of New Hampshire v. Joel Hungerford* 94-S-45 through 47; *State of New Hampshire v. John Morahan* 93-S-1734. (Hillsborough County Superior Court, May 23, 1995). [1MB]
- Stern, D. (1991) Models of memory: Wittgenstein and cognitive science. *Philosophical Psychology* 4:203-18. [AB-Z]
- Subkoviak, M. J. (1975) The use of multidimensional scaling in educational research. *Review of Educational Research* 45:387-423. [aAK]
- Sulsky, L. M. & Balzer, W. K. (1988) Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology* 73:497-506. [aAK]
- Tulving, E. (1962) Subjective organization in free recall of "unrelated" words. *Psychological Review* 69:344-54. [aAK] (1979) Memory research: What kind of progress? In: *Perspectives on memory research: Essays in honor of Uppsala University's 500th anniversary*, ed. L.-G. Nilsson. Erlbaum. [aAK] (1983) *Elements of episodic memory*. Clarendon Press. [aAK, RPF, BLS, EW]

## References/Koriat & Goldsmith: Memory metaphors

- (1985a) How many memory systems are there? *American Psychologist* 40:385-98. [aAK] (1991) Memory research is not a zero-sum game. *American Psychologist* 46:41-42. [aAK, EG]
- Tulving, E. & Pearlstone, Z. (1966) Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior* 5:381-91. [EW]
- Tulving, E. & Thomson, D. M. (1973) Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80:352-73. [aAK, LK]
- Tversky, B. (1981) Distortions in memory for maps. *Cognitive Psychology* 13:407-33. [aAK]
- Tversky, B. & Schiano, D. J. (1989) Perceptual and conceptual factors in distortions in memory for graphs and maps. *Journal of Experimental Psychology: General* 118:387-98. [aAK]
- Wagenaar, W. A. (1986) My memory: A study of autobiographical memory over six years. *Cognitive Psychology* 18:225-52. [EW]
- Wagenaar, W. A. & Boer, H. P. A. (1987) Misleading postevent information: Testing parameterized models of integration in memory. *Acta Psychologica* 66:291-306. [aAK]
- Waterman, S. & Gordon, D. (1984) A quantitative-comparative approach to analysis of distortion in mental maps. *Professional Geographer* 36:326-37. [aAK]
- Watkins, M. J. (1990) Mediationism and the obfuscation of memory. *American Psychologist* 45:328-35. [aAK]
- Watkins, M. J. & Tulving, E. (1975) Episodic memory: When recognition fails. *Journal of Experimental Psychology: General* 104:5-29. [aAK, LK]
- Waugh, N. C. & Norman, D. A. (1965) Primary memory. *Psychological Review* 72:89-104. [aAK]
- Weiner, B. (1966a) Effects of motivation on the availability and retrieval of memory traces. *Psychological Bulletin* 65:24-37. [aAK] (1966b) Motivation and memory. *Psychological Monographs: General and Applied* 80(18) Whole No. 626. [aAK]
- Weingardt, K. R., Leonesio, R. J. & Loftus, E. F. (1994) Viewing eyewitness research from a metacognitive perspective. In: *Metacognition: Knowing about knowing*, ed. J. Metcalfe & A. P. Shimamura. MIT Press. [aAK]
- Wells, G. L. & Lindsay, R. C. L. (1985) Methodological notes on the accuracy-confidence relation in eyewitness identifications. *Journal of Applied Psychology* 70:413-19. [aAK]
- Wells, G. L. & Loftus, E. E., Eds. (1984) *Eyewitness testimony: Psychological perspectives*. Cambridge University Press. [aAK]
- Wells, G. L., & Turtle (1988) What is the best way to encode faces? In: *Practical aspects of memory: Current research and issues*, vol 1, ed. M. Gruneberg, P. Morris & R. Sykes. Wiley. [aAK]
- White, K. G. (1985) Characteristics of forgetting functions in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior* 44:15-34. [KGW]
- (1991) Psychophysics of direct remembering. In: *Signal detection: Mechanisms models and applications*, ed. M. L. Commons, J. A. Nevin & M. C. Davison. Erlbaum. [KGW]
- White, K. G. & Cooney, E. B. (in press) Consequences of remembering: Independence of performance at different retention intervals. *Journal of Experimental Psychology: Animal Behavior Processes*. [KGW]
- White, R. T. (1982) Memory for personal events. *Human Learning* 1:171-83. [aAK]
- Wilkening, F. & Anderson, N. H. (1991) Representation and diagnosis of knowledge structures in development psychology. In: *Contributions to information theory: vol. 3. Developmental*, ed. N. H. Anderson. Erlbaum. [NHA]
- Winograd, E. (1988) Continuities between ecological and laboratory approaches to memory. In: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, ed. U. Neisser & E. Winograd. Cambridge University Press. [aAK] (1991) The case for both approaches to memory. Paper presented at the International Conference on Memory (July 1991), Lancaster, U.K. [LK]
- (1994) Comments on the authenticity and utility of memories. In: *The remembering self: Construction and accuracy in the self-narrative*, ed. U. Neisser & R. Fivush. Cambridge University Press. [aAK]
- Winograd, E. & Neisser, U., Eds. (1992) *Affect and accuracy in recall: Studies on "flashbulb memories."* Cambridge University Press. [aAK]
- Wittgenstein, L. (1980) *Remarks on the philosophy of psychology*, vol. 1. Blackwell. [MAC]
- Wixted, J. T. & Ebbesen, E. (1991) On the form of forgetting. *Psychological Science* 2:409-415. [KGW]
- Woodworth, R. S. (1938) *Experimental psychology*. Holt. [HPB]
- Wright, D. B. (in press) Methodological issues for naturalistic event memory research. To appear in: *Intersections in basic and applied memory research*, ed. D. Payne & F. Conrad. Erlbaum. [DBW]
- Yaniv, I. & Foster, D. P. (1990) Judgment, graininess, and categories. In: *Proceedings of the Twelfth Annual Cognitive Science Society*. Erlbaum. [aAK] (1994) Graininess of judgement under uncertainty: An accuracy informativeness trade-off. *Journal of Experimental Psychology: General* 124:424-32. [aAK]
- Yaniv, I., Yates, J. F. & Smith, J. E. K. (1991) Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin* 110:611-17. [aAK]
- Yates, J. F. (1982) External correspondence: Decompositions of the mean probability score. *Organizational Behavior and Human Performance* 30:132-56. [aAK] (1990) *Judgment and decision-making*. Prentice-Hall. [aAK]