The Relationship Between Memory and Judgment Depends on Whether the Judgment Task is Memory-Based or On-Line

Reid Hastie  
Northwestern University

Bernadette Park  
University of Colorado

Five alternative information processing models that relate memory for evidence to judgments based on the evidence are identified in the current social cognition literature: independent processing, availability, biased retrieval, biased encoding, and incongruity-biased encoding. A distinction between two types of judgment tasks, memory-based versus on-line, is introduced and is related to the five process models. In memory-based tasks where the availability model describes subjects' thinking, direct correlations between memory and judgment measures are obtained. In on-line tasks where any of the remaining four process models may apply, prediction of the memory–judgment relationship is equivocal but usually follows the independence model prediction of zero correlation.

There ought to be a relationship between memory and judgment. Our intuition tells us that we should be able to generate more arguments and information in support of a favored position than against it, that evaluations of people should be related to the amounts of good and bad information we have about them. When a person is able to remember many arguments against a belief, or to cite many good characteristics of an acquaintance, we are surprised if they endorse the belief or dislike the person. In support of intuitions like these, names have been given to the idea that memory and judgment have a simple direct relationship, including “availability,” “dominance of the given,” “salience effect,” and so forth.

However, empirical studies of the relationship between memory and judgment with subject matter as diverse as social impressions, personal attitudes, attributions of causes for behavior, evaluations of legal culpability, and a variety of probability and frequency estimates have not revealed simple relations between memory and judgment. Some relationships have been found, but strong empirical relations are rare and results are often contradictory.

Some examples seem to support the expectation of a direct relationship between memory and judgment. Tversky and Kahneman (1973) demonstrated that many judgments of numerosity were directly correlated with the “ease with which instances or associations could be brought to mind” (p. 208). In an illustrative series of experiments, they showed that judgments of the frequency of words in English text were correlated with the ease of remembering the words. Beyth-Marom and Fischhoff (1977) provided more definite evidence on the strength of the memory–judgment relationship, and Gabrieleick and Fazio (1984) demonstrated that ease of retrieval exerts a causal (not just correlational) influence on the frequency estimates.

Similar demonstrations have been provided in research studying memory–judgment relationships for more complex naturally occurring events. Lichtenstein, Slovic, Fischhoff, Layman, and Combs (1978; Combs & Slovic, 1979) obtained large correlations between judged frequencies of deaths and the frequencies of reports of causes of death in newspaper articles. For these events, reporting rates were not correlated with actuarial frequencies, supporting the conclusion that newspaper reporting rates biased memory availability, which in turn influenced the frequency estimates. Another neat everyday example is provided by Ross and Sicoly (1979) in their research on the attribution of responsibility: They found simple, direct relationships in joint tasks such as home improvement projects, writing scientific papers, and competing in team sports between subjects' judgments of responsibility for themselves and others and the subjects' ability to remember relevant evidence. These examples imply a direct relationship between memory and judgment: The more a subject can remember and the more easily it comes to mind, the higher the estimates of frequency, probability, or responsibility.

However, there is another side to the literature on memory–judgment relationships. Although researchers almost invariably expect to find the direct relationships illustrated by the first set of examples, this has not always been the case. In a classic experiment by Anderson and Hubert (1963), subjects made likability ratings of a hypothetical character on the basis of sentences of trait adjectives. In some conditions of the experiment, subjects were also asked to recall the adjectives. Anderson and Hubert noted that when the recall data were summarized as serial position curves, a recency effect was obtained; however, the likability rating data showed that early adjectives in the sequence had the greatest impact on subjects' final impressions, a primacy effect. Anderson and Hubert concluded that “the impression memory is distinct from the verbal memory for the adjectives” (p. 391).

Another frequently cited study of memory–judgment rela-
tionships is an experiment by Reyes, Thompson, and Bower (1980) in which subjects were asked to make judgments of the apparent guilt of a defendant briefly described in a drunken driving case. These experimenters manipulated information about the defendant's character and the memorability of items of evidence favoring defense or prosecution sides of the case. The authors did not find a relationship between memory and judgment at short delays; however, after a 48-hour delay a significant but low correlation (+.31) was obtained. The authors concluded that "such a low correlation suggests that judgments of apparent guilt were not made simply on the basis of the differential availability of the arguments" (p. 7).

Finally, there is a third relationship between memory and judgment that appears in some person impression formation tasks. Hastie and Kumar (1979; replicated by many others, most notably Srull, 1981) found that when subjects studied sentences describing the behavior of a hypothetical character while making judgments of the character's personality, behaviors that were incongruent with the final trait ratings were the best recalled on a subsequent memory test. This reversal of the commonly expected direct memory- judgment relationship has been explained as resulting from "special processing" accorded to the incongruent acts when they were attributed to a single character in the context of an impression formation task (Hastie, 1984). Empirical results suggest that this special processing involves the subject's effort to explain why the surprising, incongruent actions were performed by the character, a plausible subtask within the larger impression formation task.

The scope of the literature from which these examples were culled is large. We have identified more than 50 published experiments that report correlational or treatment effect findings relevant to the memory-judgment relationship in judgment tasks. The variety of judgment task domains included in these studies is remarkably diverse even within the social judgment area: impression formation involving traits, behaviors, likability, and ability judgments; occupation suitability judgments; attitude formation and change; judgments of historical events' antecedents and consequences; predictions of sports events; psychodiagnosis of clinical cases; causal attribution for individual behavior; statistical decision making; and deductive reasoning.

Theoretical Analysis

The contrast between our expectations that memory-judgment relationships should be direct and pervasive and the mixed results of research on the relationship is puzzling. We believe that a careful analysis of the cognitive processes that could produce or obscure a memory- judgment relationship can explain why the relationship appears or does not appear across experimental tasks. Several specific models for the perception, memory, and judgment processes that could produce correlations between memory and judgment measures have been presented by researchers studying social judgment. To resolve the enigma concerning when memory and judgment will be directly related, we need to clarify the alternate process models and identify the empirical conditions under which each model will apply.

Our theoretical analysis is nested in the context of an information processing approach to social cognition (Hastie & Carlston, 1980; Smith, 1984; Wyer & Srull, 1980). Briefly, we assume that when a person is presented with a judgment task, either in an experimental or natural situation, evidence information is processed by a judgment operator that performs its function to generate a conclusion on which a response is based. Traditionally, judgment researchers have described these operators as algebraic combination rules (e.g., Anderson, 1981; Edwards, 1968; Hammond, Stewart, Brehmer, & Steinmann, 1975), stochastic processes (e.g., Thomas & Hogue, 1976), or cognitive heuristics (e.g., Kelley, 1973; Tversky & Kahneman, 1974).

The judgment operator is limited by working-memory capacities constraining the complexity of elementary information processes that can be executed at any point in time. The same limits apply to the activation of memory representations and constrain the amount of evidence that can be input into the operator at any moment. Information processing researchers have identified many information formats and structures that might characterize the evidence that is input into the operator. They have also described many potential judgment operators that have been learned over a person's lifespan and are stored in long-term memory along with other cognitive skill procedures.

Five Information Processing Models

The current social cognition literature provides several examples of memory-judgment models. They can be separated into three classes according to the causal priority of memory or judgment processes: no-priority-independence, memory causes judgment, and judgment causes memory. The most populated class comprises judgment-causes-memory models (biased retrieval, biased encoding, and incongruity-biased encoding), with one example each of an independence model (two-memory hypothesis) and a memory-causes-judgment model (availability-biased judgment). Note that the establishment of a direct or indirect correlation between memory and judgment measures will not answer the question of whether memory has causal priority and causes judgment or the reverse, judgment causes memory. Furthermore, most research has not attempted to go beyond correlational analysis to resolve the ambiguity and establish causal relationships (Dellarosa & Bourne, 1984; and Gabrieleck & Fazio, 1984, are possibly the only exceptions). However, we will review the alternate theoretical models before we turn to the tough questions of empirical prediction and identifiability.

Independence. First, there is a basic model that postulates independence between judgment and memory processes, with no relationship expected between measures of memory and judgments. Norman Anderson (1981; Anderson & Hubert, 1963) has probably stated this position most clearly in his articles on the two-memory hypothesis in trait-adjective-based impression formation tasks:

The judgment was based on a memory system different from the recall... As each adjective was received, the valuation operation extracted its implications for the task at hand. Further processing, especially the integration, was performed on these implications. The verbal material itself, no longer necessary, was transferred to a verbal memory or forgotten. (Anderson, 1981, pp. 95-96)
Thus, the valuation and integration operations involved in the execution of the judgment operator (usually summarized as an algebraic weighted averaging calculation) and the encoding of evidence information into long-term memory traces occur simultaneously and independently.

**Availability.** Second is the currently popular family of availability models which assume that memory availability causes judgment. The model has many precedents, but a seminal article by Tversky and Kahneman (1973) on the "availability heuristic" is its most common source. This model can be stated as follows: (a) During the time when evidence information is available in the external environment, the subject encodes that information in working memory. The judgment is not made at this time; usually the subject is unaware that the information is relevant to a future judgment. (b) At this time only one process operates that is relevant to the later memory–judgment relationship, namely, the further encoding of the evidence information by transforming it from working-memory codes into long-term memory traces. (c) At some later point in time when a judgment is called for (by the experimenter or "spontaneously"), the subject initiates the judgment process and retrieves information from long-term memory to use as input into a judgment operator. (d) A judgment is generated and reported to the experimenter at this point on the basis of evidence retrieved from long-term memory. (e) When the experimenter tests memory, the memory retrieval process is repeated (in essentially the same fashion as it occurred to generate input for the judgment operator) and the subject responds on the memory test. A relationship is produced between judgment and memory because any tendency that the subject may have to selectively remember information will be reflected both in biased input to the judgment operator and in the biased sample of information reported on the memory test.

The availability process model we described is slightly more limited than the original Tversky and Kahneman (1973) heuristic. We exclude cases of perceptual availability and restrict the model to memory availability. We also exclude consideration of the availability of previously made inferences and judgments to focus on memory availability of relatively raw evidence information (see Lingle & Ostrom, 1979, for a different opinion). These more complex cases involving judgments based on the retrieval of other judgments and principles of perceptual salience are treated as derived from our five elementary models (see discussion below).

In the introduction to this article we gave examples of experiments by Tversky and Kahneman, Ross and Siculo, Lichtenstein et al., and others who reported correlational results consistent with the availability heuristic memory-causes-judgment hypothesis. The memory-causes-judgment process model that we sketched above is surely correct in many applications. However, one point of the present article is that the availability heuristic model has been generalized to many tasks to which it does not apply. The evidence for overgeneralization appears in many articles in which a strong correlation between memory and judgment is not obtained, although the authors often persist with the availability heuristic interpretation in their discussion sections (e.g., Fiske, Kenny, & Taylor, 1982; Reyes et al., 1980).

**Biased retrieval.** A second set of causal models assumes that judgment causes memory. The most common process model assumes that the judgment and memory-encoding processes go on at about the same time, independently, but the judgment, once completed, has the potential to bias retrieval. (a) Evidence information occurs in the environment and is encoded into a representation in working memory. (b) During its stay in working memory, the information serves as input for two independent processes: information is transformed and written into long-term memory traces and information is operated on to produce a unitary judgment conclusion. (c) When the experimenter asks the subject to report his or her judgment, the subject responds by considering the judgment conclusion that is stored in working memory or, if time has passed, in long-term memory. (d) When the experimenter tests the subject’s memory for the evidence, the subject searches long-term memory to find trace information to respond on the memory test. At this point the subject’s judgment plays a special role in that it biases access to traces stored in long-term memory or serves to edit the traces as they are output in the form of responses on the memory test. Thus, the judgment serves a selective function such that traces that “fit” the judgment are likelier to be found in the memory search or to be reported at the memory decision stage. Common names for this class of biases include “selective recall,” “confirmatory memory,” and “access-biased memory” (Leamer, 1974, 1975; Snyder & Uranowitz, 1978). The selective influence of the judgment during the retrieval phase of a memory task biases recall in a fashion that produces a correlation between the judgment (c) and the memory (d) responses.

There is some controversy in the literature about exactly what mechanism at retrieval produces the bias. Snyder and Uranowitz (1978) suggested that the search phase of retrieval was the locus of the bias, but subsequent studies by Bellezza and Bower (1981) and Clark and Woll (1981) favored an editing or response threshold mechanism to explain the relevant findings. For present purposes it will suffice to cite the Snyder and Uranowitz results as illustrating the basic retrieval bias phenomenon because either biased search or biased responses would fit the slightly more general model we are describing.

**Biased encoding.** The fourth model assumes that judgment causes memory by biasing the encoding of evidence information. (a) Evidence information enters working memory and is used as input to the judgment operator in the formation of an initial judgment. (b) The initial judgment filters subsequent evidence information by guiding information search, encoding, and comprehension in such a way that information that fits the initial judgment (or that can be interpreted to fit) has an advantage in subsequent transfer to long-term memory. (c) When the experimenter asks the subject to report his or her judgment, the subject responds by considering the judgment conclusion that is stored in working memory or, if time has passed, in long-term memory. (d) When the experimenter tests the subject’s memory for the evidence, the subject searches long-term memory to find trace information to respond on the memory test. Because the evidence stored in long-term memory was filtered through the lens of the initial judgment, memory search will locate a biased sample of information reflecting the encoding bias.

Oddly enough, although this is the most frequently mentioned judgment–memory model, there is the least support for its applicability to social judgment phenomena (Alba & Hasher, 1983; Hastie, 1981). In our view, the social psychological find-
ings that have been cited as examples of the model are more plausibly interpreted as biased-retrieval phenomena (see Taylor & Crocker, 1981, for a typical review). We must admit, however, that some examples are ambiguous and could represent either biased-encoding or biased-retrieval processes (e.g., Berman, Read, & Kenny, 1983).

Incongruity-biased encoding. Finally, there is a model which assumes that judgment causes memory during the encoding stage, but the hypothesized memory–judgment relationship is the reverse of the common filtered encoding bias: we call this the incongruity-biased-encoding model. Hastie (1980, 1984) and Srull (1981) have given this model its clearest statement in interpreting the results of a series of studies of impression formation and recall tasks in which judgment-incongruent evidence was better recalled than judgment-congruent evidence (see also Woll & Graesser, 1982, for an interpretation that emphasizes incongruity-biased-encoding processes). (a) Evidence information is encoded in working memory and is input to a judgment operator where an initial judgment is formed. (b) Later incoming information is reviewed in the context of the initial judgment, and information that is incongruent (contradictory) is given special processing that enhances its memorability (e.g., it receives more associative links to other information in working memory, or it receives "special tags" that strongly attach it to a knowledge structure that organizes current events in long-term memory). (c) When the experimenter asks the subject to report a final judgment, the subject responds by considering the judgment that is stored in working memory or in long-term memory. (d) When the experimenter tests the subject's memory for the evidence, the subject searches long-term memory and is especially likely to find incongruent information. The incongruent information advantage accrues from the rich network of associative links or the special tags that were attached to incongruent evidence during encoding. The most frequently cited examples of this model are found in the Hastie and Kumar (1979) and Srull (1981) person memory experiments cited above.

Given that information processing can plausibly follow one of several alternate models, the derivation of unequivocal predictions concerning the direction and strength of memory–judgment relationships, even at the gross correlational level, will not be a simple matter. Furthermore, it is likely that in many laboratory and naturally occurring judgment tasks more than one of the theoretical process models may apply to a single subject's performance (e.g., both encoding and retrieval biases may operate or both judgment-causes-memory and memory-causes-judgment processes may apply; Hastie, 1981). We should not be surprised that empirical results are mixed, given theoretical hypotheses to account for either direct or indirect relationships. We believe that careful thought about the task conditions that elicit each model can resolve some of the theoretical and empirical confusions.

On-Line Versus Memory-Based Judgment Tasks

The key distinction we would like to make concerns the source of inputs to the judgment operators. In essence our hypothesis is that in some tasks the judgment is necessarily memory based and the subject must perform rely on the retrieval of relatively concrete evidence from long-term memory in order to render a judgment. Under these memory-based conditions, direct relationships between memory for the evidence and the judgment will be obtained. The direct relationships would be predicted from either of the availability process models that would apply to memory-based tasks. However, these conditions are rare, both in natural and laboratory environments, and in-genuity must be exercised to design tasks that will produce simple, direct memory–judgment relationships. The alternative, more common, class of on-line judgment tasks is associated with several process models that do not all yield predictions of a direct memory–judgment relationship.

In many judgment tasks, information for the operator follows a path from the stimulus environment external to the subject into working memory and directly to the judgment operator. We call tasks of this type on-line judgment tasks because the subject is forming the judgment "on-line" as evidence information is encountered. For example, in most research on impression formation (Anderson & Hubert, 1963; Asch, 1946), adjectives are presented to subjects as the basis for an impression and they are used by the judgment operator to update the impression almost immediately after they are perceived by the subject.

A second example of an on-line, perception-based task is the abbreviated legal judgment task employed by Reyes et al. (1980), which was also described in the introduction to this article. Another well-known on-line judgment task was introduced by Ward Edwards (1968) and has been used to study judgments that can be characterized as dynamic probability revision judgments. One version of this judgment task involved subjects' inductive inferences concerning which of two bookbags was the source of a sample of poker chips. The bags differed in the composition of chips, of distinctive colors, that they contained, and a random sample of chips from the bag would be informative about which bag had been selected by the experimenter. Lopes (1982; see also Einhorn & Hogarth, 1985) proposed what we believe is the most plausible judgment operator for performance in this task. Her "anchoring and adjustment" model for opinion revision is a prototype of an on-line judgment procedure. In all of these cases, our assumption is that subjects making the judgment revise, on-line, as items of evidence are encountered. Furthermore, the evidence items are used directly, with few intervening inferences, as inputs into the judgment operator.

Our review of the five theoretical processing models implies that it will not be possible to unequivocally predict the relationship that will be obtained between memory and judgment measures when the judgment is made on-line. It should be clear that four of the five models could apply; thus, direct, indirect, or no relationship might be observed. Because previous research has not been sensitive to the possible variety of process models and relationships that might apply in on-line judgment tasks, it is difficult to definitely identify the appropriate process model in most experiments.

The judgment procedure can take a different, memory-based course, with the input to the operator coming from long-term memory into working memory rather than directly from the external environment. The clearest examples of this condition in the judgment literature appear in the Tversky and Kahneman (1973) demonstrations of availability effects. For example, when subjects were asked to estimate the frequency of occurrence of words with certain characteristics (e.g., $k$ as the first
letter), they had to rely on information that had been stored in long-term memory to make estimates of frequencies. In addition, in the introduction to the present article we described research on risk judgments by Lichtenstein et al. (1978) and responsibility judgments by Ross and Sicoly (1979) in which subjects appeared to rely on their ability to retrieve instances that were directly relevant to a judgment from long-term memory.

If our theoretical review is valid, prediction of the memory-judgment relationship is possible in memory-based tasks. Only the availability processing model is applicable, and it predicts that there will be a direct relationship between memory and judgment measures.

The on-line versus memory-based distinction has many precedents in the social cognition literature. Two series of experiments, one directed by Lingle (Lingle, Dukerich, & Ostrom, 1983; Lingle, Geva, Ostrom, Lieppe, & Baumgardner, 1979; Lingle & Ostrom, 1979) and one by Carlson (1980) simultaneously introduced the notion of memory-based judgments to the social literature. These authors studied the case in which a subject made an initial judgment and then retrieved the first judgment from memory to serve as the basis for a second judgment. Perhaps McArthur (1980) most clearly distinguished between direct, perception-based judgment and indirect, recall-mediated judgment in the same terms as our distinction between perception and memory sources of input (see also Taylor & Thompson, 1982). Finally, Sherman and his colleagues (Sherman, Zehner, Johnson, & Hirt, 1983) and Lichtenstein and Slurr (1985) have also reached the conclusion that direct memory-judgment relationships are likeliest to be found in memory-based tasks.

Why is the on-line versus memory-based task difference the key to understanding the mixed results from empirical research on the memory-judgment relationship? If we can arrange conditions to insure that a subject makes a memory-based judgment, the memory-judgment relationship should fit the availability process model, which predicts a direct relationship between evidence memorability and the judgment conclusion. However, if the judgment task is on-line, four of the five models may apply. Therefore the memory-judgment relationship can be either direct, indirect, or null, and unequivocal prediction is not possible.

Because so many conditions are likely to instigate perception-based, on-line judgments, the difficult question for experimenters on the memory-judgment relationship is how to produce memory-based judgments. Memory-based judgments are usually more effortful than on-line judgments, and people realize that memory processes are less reliable than the perceptual processes involved in on-line judgments. Hence, subjects make on-line judgments when they believe that a judgment is likely to be required at a later point in time.

What are the conditions that will produce memory-based judgments? Probably the most reliable method to produce memory-based judgments is to surprise subjects with a novel judgment that is unlikely to be preceded by a relevant on-line judgment. Everyday examples of memory-based “surprise judgments” abound. A university professor returns from a professional meeting and learns that her department has a position open. Relatively effortful memory-based judgments are made of prospective candidates who had been encountered at the meeting. Similarly, a student might discover that he is looking for a roommate and then review acquaintances on the basis of memory to make preliminary judgments of compatibility. In these examples it is obvious that if the judgments had been anticipated before the evidence was encountered, the judgments would have been on-line.

On-Line Spontaneous Judgments

Normally people make many judgments spontaneously, without waiting for an instruction from one of life’s experimenters. We often make spontaneous inferences about other people, particularly in the first encounter situations so frequently simulated in laboratory and field research on impression formation, stereotyping, attribution, moral evaluation, and persuasion. Essentially we assume that most judgments studied by social psychologists are perception-based, on-line processes. Even the reader who believes that people are mindless drones in most natural social interactions (Langer, Blank, & Chanowitz, 1978) will admit that psychology experiments on social judgment are an exception to the norm of mental inactivity. In these experiments judgments of interest to the experimenter are almost always being made on-line by subjects because of the nature of the situation (e.g., “This is a hypothetical acquaintance situation”; “We are interested in your first impressions”; “Suppose that you are a personnel officer”); the task (e.g., “You will make judgments of the likability of 12 individuals”); or specific instructions (e.g., “Attempt to evaluate the relative degrees of responsibility . . .”).

Recently social psychologists have conducted research that identifies some of the major types of social judgments that will tend to be made spontaneously, on-line, “in the absence of an investigator’s instructions” (Winter & Uleman, 1984, p. 237). In accord with our above assumption, trait (Winter & Uleman, 1984) and causal (Weiner, 1985) inferences are likely to be made spontaneously. Furthermore, several additional classes of factors have been found to elicit spontaneous social judgments.

First are a set of individual difference factors that are frequently referred to as “schematism” or “chronicity” in the social psychology literature. For example, Markus (1977) demonstrated that people exhibit differential tendencies to make judgments of other people along certain dimensions or with reference to certain attributes. Some people appear to be constantly concerned with the masculinity–femininity of others, some people with the dependence–independence of others, some people with the intelligence of others, and so forth. These individual differences lead subjects to spontaneously make inferences about almost everybody they meet, although the specific dimensions will differ across subjects (see also Bargh, 1982; Higgins, King, & Mavin, 1982).

Second, Hastie (1980, 1984) has speculated that unexpected events will tend to elicit certain types of social judgments, particularly causal attribution judgments. Third, it seems that other stimulus conditions can also invoke spontaneous judgments: stimulus properties that are strikingly and distinctively associated with certain social categories or activities. For example, it is difficult to look at Arnold Schwarzenegger without making a judgment of his athletic prowess. Fourth, there is considerable speculation about the relationships between social
goals (e.g., to form a clear impression of another person, to induce another to like oneself, to persuade another to comply with one's desires, to inform another about one's [self-conceived] identity, etc.) and judgment. Weiner (1985) concluded that a goal of "mastery" is the primary instigating condition for causal attribution judgments. Similarly, Jones and Pittman (1982) hypothesized that many types of judgments would be made to facilitate "strategic self presentation" in the service of a motive to exert power over others. More specific social goals have been identified by other theorists (e.g., Hamilton, Katz, & Leirer, 1980; Hoffman, Mischel, & Mazzie, 1981; Swann, 1984), but little empirical research on the goal-judgment link has been conducted.

In fact, there is a further complication: Even when subjects have not spontaneously made the judgment requested by an experimenter, they seem reluctant to consult long-term memory for evidence on which to base a novel judgment. Rather, they appear to prefer to make a new judgment on the basis of earlier judgments and inferences without retrieving specific evidence from memory (Lingle, Dukerich, & Ostrom, 1983; Lingle & Ostrom, 1979). Thus, true memory-based judgments may be rare because so many judgments are made on-line (spontaneously) and because when a new judgment must be made in the absence of perceptually available evidence, subjects rely on previous judgments rather than remembered evidence.

To summarize, we started with the empirical enigma of the rarity of direct memory-judgment relationships in the results of experiments designed to obtain them. Five models for the information processes that might underlie memory-judgment relationships were abstracted from the social cognition literature: the two-memory independence hypothesis, availability, retrieval bias, encoding bias, and incongruity-biased encoding. The independence model predicts no correlation between memory and judgment measures, the incongruity-biased-encoding model predicts a negative correlation, and the remaining models predict a positive correlation. A distinction was introduced between on-line and memory-based judgment tasks, and we hypothesized that the availability model was the only candidate to apply in memory-based tasks. However, most laboratory and real-world judgments probably occur under on-line task conditions where there is uncertainty about which of the remaining four process models apply, and thus the memory-judgment relationship is equivocal.

The bottom line of our analysis is that to guarantee a direct relationship between memory and judgment measures, we must study memory-based judgment tasks that occur (a) when no previous relevant judgment has been made and (b) when the subject is motivated to use evidence retrieved from memory as input at the time of judgment. In the next section we will summarize the results of four experiments that vary task conditions to produce the predicted direct relationship between memory and judgment.

Empirical Research

The most convincing argument we can make in support of our conceptual analysis is to conduct research showing that empirical phenomena fit our predictions. Results from four experiments that investigated the memory-judgment relationship are summarized in Table 1. In each experiment, task conditions were arranged to force subjects to adopt a memory-based judgment strategy, where we would predict substantial correlations between memory and judgment measures. In the last three experiments we added task conditions that would either encourage on-line judgment strategies or would allow subjects to make the final judgment based on a previous spontaneous judgment retrieved from memory. In these conditions we did not predict correlations between memory and judgment. We will briefly sketch the methods and results from each experiment.

In Experiment 1 subjects were asked to make a judgment of a man's suitability for a job as a computer programmer after hearing a 5-min conversation between two men. Half of the subjects were told of the judgment before hearing the conversation (on-line task condition), and half learned of the judgment only after the conversation (memory-based task condition). Pretesting assured us that the job suitability judgment would not be made spontaneously by subjects in the memory-based task. After the tape-recorded conversation, all subjects made the job suitability judgment (on a 10-point rating scale) and recalled as much information from the conversation as they could. Order of the rating and recall tasks was varied, but it did not affect the results. Experimental instructions, identical except for timing in on-line and memory-based conditions, stressed accuracy in judgment and completeness of recall to counter any tendencies that subjects might have to follow an unspoken experimental demand for agreement between recall and judgment.

Several indices of the strength of recall of items favoring or opposing the job candidate were calculated, and all supported the same conclusions. The measure we will report is a ratio (calculated for each subject) of the number of items recalled favoring suitability divided by the number of items recalled favoring and opposing suitability. A ratio of .50 corresponds to recall of an equal number of items favoring and opposing, greater than .50 indicates more favoring items recalled, and less than .50 indicates more opposing items recalled. A similar ratio was calculated on the basis of only the first five items recalled, on the assumption that in a memory-based judgment process, primacy in retrieval would determine item impact on the judgment (i.e., the subject would probably not recall more than five items before rendering one of our experimental judgments, and the earliest items retrieved would have the greatest impact on the judgment).

The measures of the relationship between memory and judgment reported in Table 1 are simple correlation coefficients calculated across subjects in each judgment condition (on-line vs. memory-based) between the subjects' recall ratios and judgments of job suitability. In all cases where a relationship was discernible, the function relating judgment and recall measures appeared to be linear, based on an inspection of the residuals from a fitted line. Our prediction was confirmed: The correlations between memory and judgment measures are substantial in the memory-based task (+.46 and +.42, ps < .05) but not in the on-line task (-.14 and +.14, ns).

Experiment 2 provides a replication of these results for the

---

1 A detailed report, which has been submitted for publication, is available from the authors.
Table 1
Correlations Between Memory and Judgment Measures From Four Experiments

<table>
<thead>
<tr>
<th>Judgment</th>
<th>Total recall</th>
<th></th>
<th>First five recall</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On-line</td>
<td>Memory</td>
<td>On-line</td>
<td>Memory</td>
</tr>
<tr>
<td></td>
<td>( r )</td>
<td>( n )</td>
<td>( r )</td>
<td>( n )</td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job suitability</td>
<td>-.14</td>
<td>35</td>
<td>.46</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>.14</td>
<td>35</td>
<td>.42</td>
<td>33</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job suitability (nonsens.)</td>
<td>.24</td>
<td>20</td>
<td>.39</td>
<td>23</td>
</tr>
<tr>
<td>Gender (sens.)</td>
<td>.37</td>
<td>20</td>
<td>.20</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>.21</td>
<td>20</td>
<td>.44</td>
<td>23</td>
</tr>
<tr>
<td>Experiment 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise (nonsens.)</td>
<td>.09</td>
<td>64</td>
<td>.56</td>
<td>65</td>
</tr>
<tr>
<td>Sociability (sens.)</td>
<td>.11</td>
<td>65</td>
<td>-.09</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>.26</td>
<td>20</td>
<td>.47</td>
<td>22</td>
</tr>
<tr>
<td>Experiment 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>.15</td>
<td>53</td>
<td>.67</td>
<td>39</td>
</tr>
<tr>
<td>Friendliness</td>
<td>.29</td>
<td>53</td>
<td>.73</td>
<td>41</td>
</tr>
<tr>
<td>Likability</td>
<td>.10</td>
<td>53</td>
<td>.56</td>
<td>43</td>
</tr>
</tbody>
</table>

The results from Experiment 1 are replicated in the job judgment condition of Experiment 2 (\( r_s \) of .39 and .44, \( p_s < .05 \), for memory-based judgments; \( r_s \) of .24 and .21, \( n_s \), for on-line judgments). Results for the gender judgment are not exactly in line with our predictions, although they do not strictly contradict our hypothesis. Essentially there is no clear difference between on-line and memory-based tasks in the strength of the memory–judgment correlations. The total recall ratio analysis suggests that the relationship in the on-line condition is stronger (.37 vs .20), whereas the first five recall ratio yields the reverse order (.26 vs .47). However, although our theoretical analysis left us uncertain about what, if any, relationship to expect in spontaneous conditions, we were surprised to obtain substantial recall ratio–judgment correlations.

Experiment 3 shifts to a new content domain and to judgments of numerosity rather than suitability or category membership. Subjects listened to a tape-recorded monologue attributed to another undergraduate describing his activities during the past week. Two judgment conditions were used: (a) one selected to be spontaneous, a judgment of the frequency with which the speaker interacted with women (sociability and gender were the judgments most commonly cited as spontaneous by pretest subjects) and (b) one selected to be nonsensuous, a judgment of the frequency with which the speaker engaged in cardiovascular exercise (none of the pretest subjects cited this type of judgment as occurring spontaneously). As before, half of the subjects were given instructions for the judgment task before the tape (on-line) and half only after the tape (memory). In this experiment every subject received one judgment before the tape and both after, so that each subject provided data for one on-line task (sociability or exercise) and one memory-based task (exercise or sociability). Note that because we have identified the sociability judgment as spontaneous, we do not expect to obtain differences between on-line and memory-based subjects for this judgment.

The results from Experiment 3 confirm our predictions for the exercise judgment; a relationship between memory and judgment measures was observed in the memory-based but not in the on-line task conditions (\( r_s \) of .56, \( p < .05 \), and .09, \( n_s \)). The results for the sociability judgment were also in line with our prediction; no relationship between memory and judgment measures in either memory-based or on-line tasks (\( r_s \) of .11 and -.09, both \( n_s \)).

Experiment 4 provides new materials, judgment tasks, and a new operationalization of the on-line versus memory-based task distinction. Subjects listened to a tape-recorded list of 30 sentences describing a person engaged in everyday actions that had implications for the person’s intelligence and friendliness (see Hastie & Kumar, 1979, for a description of the materials). Previous research (Winter & Uleman, 1984) and our pretest ratings suggested that personality and ability judgments would usually be made spontaneously by our undergraduate subjects. Thus, the problem of operationalizing the on-line versus mem-
ory-based difference becomes the problem of turning a spontaneously made (on-line) judgment into a memory-based judgment. Note that in the previous experiments our problem was to turn a rarely made nonspontaneous judgment into an on-line judgment. In those studies, we used experimenter instructions before presenting evidence to make the nonspontaneous judgment on-line. In Experiment 4 we attempted to interfere with the subjects' tendency to make personality and ability judgments spontaneously by presenting a novel judgment task that we expected would divert the subjects' mental processes from their normal spontaneous judgments. The task we used required subjects to make a judgment of the grammaticality of each sentence as it was presented on the tape recorder. We believed that making the grammaticality judgment would prevent the subjects from spontaneously integrating information from the sentences into impressions of the person's personality, abilities, or likability. Thus, the interfering grammaticality judgments would make impression judgments, after the tape recording, memory based. The on-line task was established by instructing subjects that they would make personality judgments at the end of the list of sentences and by requiring a likability judgment of the person after each sentence was presented.

Ratings of intelligence, friendliness, and likability from Experiment 4 subjects revealed the predicted pattern: substantial recall ratio—judgment correlations in the memory-based task (rs of .67, .63, and .56, all ps < .05) but not in the on-line (grammaticality) task (rs of .15, .29, and .10, ns). We should note that differences in the variances associated with the variables entered into any of these correlational analyses do not account for differences in the magnitudes of the coefficients. For example, there was no tendency for variances to be lower in the online task conditions in comparison with memory-based conditions for either recall ratio or judgment measures.

Results from some secondary analyses also fit the theoretical account we have advanced to explain the on-line versus memory-based task differences. Evidence items in Experiments 1 and 2 were separated into judgment-relevant and judgment-irrelevant sets to explore the effects of task (on-line versus memory-based) on the relative recall of these two types of items. One prediction from our initial theoretical account is that relevant and irrelevant items will receive differential attention and encoding in on-line tasks, where relevant items should be well remembered in comparison with irrelevant items. However, under memory-based task conditions there would be no reason to expect differential treatment of the two sets of items until after the judgment task is presented. Thus, there should be little or no difference in the recall of relevant and irrelevant items in memory-based tasks. The pattern of means of proportion recalled measures predicted by this hypothesis was observed in the results from Experiments 1 and 2: distinctively high recall of judgment-relevant items in the on-line task and no advantage in recall of relevant items in the memory-based task (all Task × Item Type interactions significant at the .05 level, means in Table 2).

One series of analyses, concerned with subjects' tendency to interpret evidence in a manner biased to fit their final judgments, did not find support for biased interpretation effects. In several of the experiments, subjects were asked, after making their judgments and recalling the evidence, to review each evidence item and to rate its implication for the judgment. We had thought that these ratings might be correlated with the final judgments (e.g., a subject who judged the job candidate as unsuited would also rate the evidence items as more strongly implying unsuitability than a subject who judged the candidate to be suited). We also speculated that the biased interpretation effect would be more pronounced in on-line tasks in comparison with memory-based tasks. However, our results did not show significant or consistent correlations between final judgments and postjudgment item ratings under any of our experimental conditions. (Of course, these results may merely indicate that our materials were relatively unambiguous and did not permit much variance in individual interpretations.)

### A Recipe for Memory–Judgment Correlations?

At times our research program has seemed like a quest for conditions sufficient to produce a significant correlation between recall ratio and judgment measures. Our theoretical clue was to seek task conditions in which subjects would be forced to rely on evidence previously stored in long-term memory and where no previously made judgment (also stored in memory) was available on which to base the judgment. In fact, our quest has been mostly successful; we did find several experimental tasks in which sizable memory–judgment relationships were reliably obtained.

The goal of our analysis is to find useful fundamental categories that classify the subject's cognitive processing strategies (five models) and to find judgment tasks (two tasks) that can serve as the building blocks to describe behavior in more complex tasks. Predictive power in the system comes from the links we have hypothesized between process models and tasks. However, we should attach a warning that several of our attempts to obtain memory–judgment correlations were not successful. Most of these failures occurred early in our research program, before we realized the importance of preventing subjects from making spontaneous on-line judgments relevant to the final judgment (in memory-based tasks). The best advice we can give researchers who extend our methods to new domains is to carefully study the types of judgments that subjects make spontaneously in their experimental tasks and then to insure that these judgments are not related to the final judgment of interest to the experimenter.

### Table 2

<table>
<thead>
<tr>
<th>Judgment</th>
<th>On-line Relevant</th>
<th>On-line Irrelevant</th>
<th>Memory-based Relevant</th>
<th>Memory-based Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>.36</td>
<td>.26</td>
<td>.26</td>
<td>.28</td>
</tr>
<tr>
<td>Job suitability judgment</td>
<td>.40</td>
<td>.34</td>
<td>.18</td>
<td>.33</td>
</tr>
<tr>
<td>Gender judgment</td>
<td>.25</td>
<td>.12</td>
<td>.17</td>
<td>.15</td>
</tr>
</tbody>
</table>
This warning is not simply a methodological note. The implication is that most judgments, laboratory and real world, do not take the simple forms of the availability and independence models we identified at the start of our analysis. Rather, conditions that invoke the biased-retrieval, biased-encoding, and incongruity-biased-encoding models associated with on-line tasks are quite common, making predictions of memory–judgment relationships difficult or impossible. We would go a step further and warn that most important judgments occur in an extended time frame and do not depend on the application of a simple operator like those identified in the algebraic, stochastic, and heuristic traditions that dominate the psychological literature. For example, even in the present experiments (2 and 3), talk-aloud reports suggest that our spontaneous judgment conditions (gender and sociability) are treated as both on-line and memory-based tasks by the subjects. They both induce an initial impression on-line, spontaneously, and rejudge in a memory-based fashion when the experimenter explicitly requests a judgment at the end of the experiment. Of course, even this complication seems trivial when compared to the duration and complexity of real-world legal (Pennington & Hastie, 1981); medical (Elstein, Shulman, & Sprafka, 1978); or diplomatic (Axelrod, 1976; Jervis, 1976) judgments. Nonetheless, simple models such as the five we begin with here constitute the proper starting point for a reduction of the complex cases to their tractable theoretical components.

Conclusion

In this article we have made the case for a distinction between two types of judgment tasks: on-line and memory-based. In support of the distinction we outlined a theoretical, information processing framework for memory–judgment relationships and identified five alternate models for processing that differed in the causal relations among memory and judgment events. The alternate models were linked to the on-line versus memory-based task distinction. The most persuasive argument for the distinction was the predicted pattern of memory–judgment relationships that we found in the results of four experiments in which on-line and memory-based tasks were presented to subjects in several social judgment domains.

Traditional theories of human judgment do not provide useful concepts to describe memory–judgment relationships. For example, cognitive algebra approaches to judgment have yielded considerable progress in our understanding of processes but have not illuminated memory–judgment relationships (Anderson, 1981). Linear statistical models of judgment policies are also mute on the memory–judgment relationship (Hammond & Summers, 1965). Similarly, approaches based on normative mathematical models would not distinguish between the memory-based and on-line task conditions that affected our subjects' behavior (Edwards, 1968).

The judgment heuristics approach was one inspiration for our information processing analysis (Tversky & Kahneman, 1974). We believe that the general form of the judgment heuristics approach provides the best current characterization of the judgment information processing system. However, the description of judgment operators, memory processes, and memory structures is extremely abstract, and specific hypotheses about the task conditions under which availability heuristics will be activated are not included in recent statements of the heuristics and biases approach (Kahneman, Slovic, & Tversky, 1982). This makes the formulation too abstract to yield a priori, testable predictions or even to describe many phenomena in the judgment literature.

The on-line versus memory-based distinction is based on differences in the sources of the information that is entered as input to the hypothetical judgment operator. Research in our laboratory suggests that the two types of tasks may affect performance in ways that are not obvious from the memory–judgment results reported here. In our experiments, talk-aloud protocols suggest that our subjects are using judgment operators that can best be described (within the larger information processing system) as algebraic combination rules similar to the weighted average or weighted sum principles identified by Anderson (1981), Hammond (1955; Hammond & Summers, 1965), and their colleagues. We believe that our subjects' operators are closest to the anchor and adjust procedures proposed by Lopes (1982) and Einhorn and Hogarth (1985) in recent extensions of the algebraic modeling approach. However, subjects do not adopt the anchor and adjust operator as frequently in memory-based tasks as in on-line tasks. Thus, there is a suggestion that the heart of the judgment process, the operator, is affected by on-line versus memory-based task conditions. A second, suggestive observation is that when evidence information is unbalanced to favor one side of the judgment, the final judgment will be both more polarized and held with more confidence in on-line as compared to memory-based tasks.

For the moment, the on-line versus memory-based distinction has made one important contribution to our understanding of human judgment. It is a mistake to look for simple memory–judgment relationships in on-line judgment tasks. Memory and judgment will be directly related, though, when the judgment is based directly on the retrieval of evidence information in memory-based judgment tasks.

References


Leamer, E. E. (1975). “Explaining your results” as access-biased mem


Ross, M., & Sicozy (1979). Egocentric biases in availability and attribu


Received July 12, 1985
Revision received October 18, 1985

Instructions to Authors

Authors should prepare manuscripts according to the Publication Manual of the American Psychological Association (3rd ed.). All manuscripts must include an abstract of 75–100 words typed on a separate sheet of paper. Typing instructions (all copy must be double-spaced) and instructions on preparing tables, figures, references, metrics, and abstracts appear in the Manual. Also, all manuscripts are subject to editing for sexist language.

APA policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more journals. APA policy also prohibits duplicate publication, that is, publication of a manuscript that has already been published in whole or in substantial part in another journal. Authors of manuscripts submitted to APA journals are expected to have available their raw data throughout the editorial review process and for at least 5 years after the date of publication.

Authors will be required to state in their initial submission letter or sign a statement that they have complied with APA ethical standards in the treatment of their sample, human or animal. (A copy of the APA Ethical Principles may be obtained from the APA Ethics Office, 1200 17th Street, N.W., Washington, DC 20036.)

Anonymous reviews are optional, and authors who wish anonymous reviews must specifically request them when submitting their manuscripts. Each copy of a manuscript to be anonymously reviewed should include a separate title page with authors' names and affiliations, and these should not appear anywhere else on the manuscript. Footnotes that identify the authors should be typed on a separate page. Authors should make every effort to see that the manuscript itself contains no clues to their identities.

Manuscripts should be submitted in quadruplicate. All copies should be clear, readable, and on paper of good quality. A dot matrix or unusual typeface is acceptable only if it is clear and legible. Dittos and mimeographed copies will not be considered. Authors should keep a copy of the manuscript to guard against loss. Mail manuscripts to the Editor, Martin L. Hoffman, Psychological Review, Department of Psychology, New York University, 6 Washington Place, 4th floor, New York, New York 10003, according to the instructions provided above.