

INTERPERSONAL REALITY MONITORING: JUDGING THE SOURCES OF OTHER PEOPLE'S MEMORIES

MARCIA K. JOHNSON, JULIE G. BUSH, AND KAREN J. MITCHELL
Princeton University

Participants judged the believability of simulated accounts varying in perceptual and emotional detail. In Experiment 1, both younger and older adults' tendency to believe that an account described an actually experienced event increased as either type of detail was added. In Experiment 2, while younger adults in a low suspicion condition judged the accounts with added details as more likely to be of actually-experienced events than the impoverished accounts, those given instructions designed to induce suspicion about the speakers' honesty found the more detailed accounts less believable. In Experiment 3, for both younger and older adults, both types of detail again increased believability ratings under low-suspicion conditions but did not affect ratings under high-suspicion conditions. In addition, there were systematic differences in the types of details the high- and low-suspicion participants reported using to make their judgments. Results are discussed in relation to the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) and reality monitoring and credibility judgments.

We hear many autobiographical tales in our everyday conversations. Only rarely do we doubt the honesty of our friends and acquaintances; usually we accept their accounts at face value (e.g., Grice, 1967/1989). But there are many reasons why someone might inadvertently or deliberately tell a tale that deviates from the truth. Friends might fail to remember a piece of information, or might remember it inaccurately. They might confuse dreams or thoughts about what could have been

This research was supported by National Institute on Aging grant AG09253 to Marcia K. Johnson, and funds provided to Julie Bush by the Psychology Department and Dean of the College of Princeton University.

The authors express their appreciation to Amy Danzer, who conducted the initial pilot study in the series, Steve Weatherall, Tina Loose, and Christine Voegeli, who helped test participants, and Robert Belli, Jonathan Schooler, and Paula Hertel, who commented on an earlier draft of this article.

Address correspondence to Marcia K. Johnson, Department of Psychology, Princeton University, Princeton, NJ 08544-1010; or, e-mail: mkj@clarity.princeton.edu.

said with a conversation that really happened. They might incorporate stories heard from others into their own autobiographical narratives. Or our friends might be lying to us about how things happened, if only to protect us from hurt feelings.

Reality monitoring is the process of attributing one's memories to external or internal origins (Johnson & Raye, 1981). Externally-derived memories originate in perceptual experience. For example, you may remember how a car looked as it ran a red light or how frightened you felt as it struck your car. Internally derived memories originate in self-generated thoughts and imagination (i.e., reflection; Johnson, 1983). For example, you may remember another occasion when you thought a car might hit you and imagined the consequences, or remember the events of a dream about an auto accident. The source monitoring framework proposes that people make attributions about the origin of such memories using heuristic processes which assess qualities of a mental experience (familiarity, perceptual detail, affect, etc.), and more systematic processes that evaluate consistency and plausibility and retrieve additional information which might help specify or rule out a given source, etc. (e.g., Johnson, 1988, 1997; Johnson et al., 1993; Johnson & Raye, 1981, *in press*).

Research on reality monitoring has generally supported this proposition, showing that memories for perceived and imagined events differ on average in their qualitative characteristics, and that people use their knowledge and beliefs about the general differences between the characteristics of classes of memories to judge the origins of particular memories (e.g., Johnson, 1988, 1997; Johnson et al., 1993). Memories for perceived events, when compared with those for reflectively-generated events, on average tend to contain more perceptual and contextual information, such as information about how things looked or where things happened (e.g., Johnson, Foley, Suengas, & Raye, 1988). They also tend to be richer in information about feelings at the time of the event (e.g., Hashtroudi, Johnson, & Chrosniak, 1990; Suengas & Johnson, 1988). In contrast, memories for imagined events tend to contain more information about the cognitive activities engaged during the event, such as processes that accompany seeing an object, imagining the sight of an object, or remembering an earlier experience involving an object (e.g., Durso & Johnson, 1980; Finke, Johnson, & Shyi, 1988). Such differences have been found in subjective ratings of the characteristics of one's memories for perceived or imagined events, as well as in experimenter ratings of free recall reports of perceived and imagined events (Conway & Dewhurst, 1995; Hashtroudi et al., 1990; Johnson et al., 1988; Schooler, Clark, & Loftus, 1988; Schooler, Gerhard, & Loftus, 1986; Suengas & Johnson, 1988; see also Johnson & Suengas, 1989).

One way, then, to evaluate the source of a memory, is to assess whether its characteristics are more typical of reflective or perceptual origins. However, applying this heuristic can sometimes lead to errors. Because distributions of actual and imagined events overlap for any given memory characteristic (e.g., some memories for imagined events have more perceptual detail than some memories for actual events), confusions among the sources of memories (i.e., source monitoring errors or source misattributions) may occur (e.g., Johnson, 1988, 1991; Johnson et al., 1993; Johnson & Raye, 1981, 1998, in press; Mitchell & Johnson, in press; Ross, 1997; Zaragoza, Lane, Ackil, & Chambers, 1997). Regardless of this possibility for error, a memory's characteristics provide critical cues about its source.

Just as the source monitoring framework has guided research regarding people's source monitoring of their own memories (see Johnson et al., 1993; Mitchell & Johnson, in press, for reviews), it can be used to investigate interpersonal reality monitoring—how people judge whether other people's memories have been perceptually experienced or imagined (e.g., Johnson & Suengas, 1989; Keogh & Markham, 1998; Schooler et al., 1986; see also Stern & Dunning, 1994). For example, if people evaluate the reality of other people's memory reports using criteria similar to those used for their own memories, they should find reports with greater amounts of perceptual, contextual, and emotional information more likely to accurately reflect perceived experiences.

Consistent with this idea, studies using protocols of remembered events show that participants' belief that these reports concerned actually-experienced events increased as a function of the amount of perceptual and contextual detail (Johnson & Suengas, 1989; Keogh & Markham, 1998; Schooler et al., 1986, 1988). Such studies involve collecting reports from people who have either perceptually experienced or reflectively generated events they described, and asking other participants to make judgments about the reports. For example, participant "judges" in Johnson and Suengas' (1989) study were asked to decide whether descriptions of events were given by earlier participants who had actually participated in the target events, or by earlier participants who had only imagined participating in the events. Descriptions judged to be actually experienced events tended to contain more perceptual and contextual information.

Although they give us a clue about factors involved in judging the origin of other people's reports, studies using real accounts cannot control for differences among reports on dimensions other than those under investigation (e.g., perceptual detail). For example, Schooler et al. (1986) asked participants to judge whether descriptions of a traffic sign were given earlier by people who had actually seen a slide of the sign,

or by people who had not seen the sign but had received post event information implying its presence. The descriptions obtained from participants who actually had seen the sign tended to be higher in perceptual content, but also shorter in length (but see Keogh & Markham, 1998) and with fewer verbal hedges. Thus, designs using real reports may not permit us to examine the specific contribution of individual qualities to, nor their relative impact on, believability judgments because the reports usually vary on other dimensions as well.

Therefore, to study more systematically the effects of various types of content details on people's judgments regarding the origin of other people's memories, the present experiments used simulated event accounts. Because the contents of the accounts were under experimenter control—that is, the accounts varied only on the variables of interest—any differences in judgments about the reports could be attributed directly to our manipulation of detail type. Accordingly, we composed a set of accounts describing commonplace experiences, such as visiting a doctor's office or studying for an exam. The accounts were relatively free from emotional or perceptual details. For each account, we also created a set of perceptual and emotional details that were inserted into the base account for some conditions. The details were designed to be incidental to the events described; they did not play a causal role and their omission in the base story version did not affect the comprehensibility of the accounts (see Appendix A for an example). We could thus compare, across participants, the impact that particular types of detail had on believability ratings of the event narrative.

Related work has been conducted in the eyewitness memory domain using a mock trial paradigm (e.g., Bell & Loftus, 1988, 1989). In one study, for example, Bell and Loftus presented descriptions of simulated cases in which defense testimony and prosecution testimony each included uninvolved eyewitnesses who testified that the defendant had either been incorrectly identified or correctly identified. A peripheral perceptual detail within the testimony of each witness was either presented in enriched form ("a box of Milk Duds and a can of Diet Pepsi") or impoverished form ("a few store items"). Results were consistent with predictions than can be derived from Johnson and Raye's (1981) reality monitoring model, in that verdict judgments and perceived eyewitness credibility were swayed in the direction of more perceptually detailed witnesses. It may be, however, that such minor details have an impact only when people are forced to choose among competing reports (cf., Bell & Jones, 1994).

In Experiment 1, we investigated whether adding perceptual or emotional detail influences believability of reports not contradicted by other reports. Furthermore, in Experiments 2 and 3, we explored the potential

interplay between participants' preconceptions about the speakers' intentions and the amount of reported detail in the accounts.

EXPERIMENT 1

In Experiment 1, we were interested in creating a situation in which the process of making judgments about other people's memories would parallel reality monitoring one's own memories. When we remember an event, there may be some question about whether the event was perceptually experienced as remembered, but the issue of whether one is lying to oneself rarely arises. Therefore, we wanted participants to evaluate accounts they believed were reported honestly and without deception. We were interested in whether participants believed that honest recountings of truly autobiographical experiences would differ from honest retellings of events not directly experienced by the speaker.

As a cover story we told participants that, in the context of a different experiment on empathy, reports had been obtained from pairs of close friends asked to recount recent memories in response to simple cues. The accounts supposedly had been generated in one of two ways. Some speakers had allegedly been asked to recount events from their own personal experiences. Others had allegedly been required to retell memories reported to them in the past by their close friends, with one constraint: they were to speak in the first person, as if they had actually experienced the events described. Note that, according to the cover story, all accounts had been honestly reported; the two sets of accounts supposedly acquired via the "empathy experiment" instructions differed only in whether or not they were truly autobiographical reports. We assessed believability by asking participants to judge whether or not they believed the events to have been actually experienced by the speaker, and to indicate their confidence in their ratings.

Our primary interest was the impact of the type of detail included in these reports on participants' believability ratings. To this end, the detail content of the accounts was manipulated by systematically varying the amount of perceptual and emotional details in each account across participants. Both emotional and perceptual details ranged from the usual (e.g., a car "with its radio on really loud") to the more unusual (e.g., a car "with a beat-up old canoe on top").

The generality of the findings was explored by including both younger and older adults as participants. Previous research on reality monitoring suggests that older adults might weight some memory characteristics—such as those regarding feelings—differently than younger adults when assessing whether their own memories are of perceived or imagined events (Johnson & Multhaup, 1992, based on Hashtroudi et al., 1990). If

older adults weight emotional information more heavily than do younger adults when judging whether a personal memory was perceptually experienced, we might expect them similarly to weight emotional information more heavily when evaluating reported memories of other people. We also investigated whether the same pattern would hold when believability judgments were made immediately after exposure to a report (i.e., "on-line"—after reading each story) and after a brief delay (i.e., from memory after reading all of the stories).

In sum, both older and younger participants in Experiment 1 read first-person accounts ostensibly given earlier by participants either reporting something experienced themselves or something actually experienced by a friend. Across participants, the accounts varied in the amount of perceptual and emotional detail. Either immediately after reading each account (immediate group) or only after reading all accounts (delay group), participants were asked to decide whether or not each account was of an event actually experienced by the speaker, and to rate their confidence.

METHOD

SUBJECTS AND DESIGN

Sixty-four Princeton undergraduates were paid participants in the experiment. Sixty-four healthy older adults, aged 61 to 81, from communities near Princeton, NJ, also served as paid participants.

Retention interval (i.e., believability ratings given immediately after reading each story vs. after all of the stories) was a between subjects factor, with equal numbers of younger and older adults randomly assigned to each condition. Type of detail (i.e., no detail added, only perceptual detail added, only emotional detail added, and both perceptual and emotional detail added) was manipulated within subjects.

MATERIALS AND PROCEDURE

Younger participants were tested in small groups of between two and 10 people. Except for one couple tested as a pair, older adults were tested individually. As a ruse to accommodate the delay manipulation, all participants were initially instructed that this study concerned impressions, and that they should try to form impressions of speakers based on the accounts they told. They were informed that the speakers were college-aged and that the accounts had been "edited to eliminate verbal fillers, hesitations, and repetitions." Participants expected to answer questions based upon their impressions.

Participants in the immediate group then received printed instructions regarding the believability-rating task (for older adults these instructions were also read aloud and discussed by the experimenter). As discussed above, participants were told that pairs of close friends had been invited to the lab for a different experiment and asked to recount true personal stories based on the experiences of either the speaker or of the friend. Participants were asked to try to distinguish, "based on [their] impressions," between two kinds of accounts: (1) accounts told by people directly involved in the events described; and (2) accounts told by people who only heard about the events from their friends, but did not experience the events themselves. Believability was rated on an 8-point scale with anchors at 1 (confident the event was not experienced by the speaker) and 8 (confident the event was experienced by the speaker). Participants were given two different opportunities to indicate that they were guessing because 4 was marked as "guessing, not experienced by speaker" and 5 was marked as "guessing, experienced by speaker." At their own pace, participants worked through a packet containing nine accounts, as described below, sequentially reading each account and rating its believability.

The procedure was somewhat different for the delay group. After receiving the initial impression-formation instructions, participants read through all accounts in their packets. Only after reading the nine accounts were they given the believability-rating instructions, as described above. Again, at their own pace they then worked through the remaining pages of their packet, each of which included an account title (e.g., Doctor's Office) and a rating scale. Stories were rated in the original order of presentation. Thus, in the delay condition, accounts were rated for believability several minutes after they had been read.

For both the immediate and the delay groups, each packet contained nine titled accounts. An unscored "buffer" story occurred in the first position for all participants to reduce primacy effects and to fully familiarize them with the judgment scale before they evaluated the eight critical accounts. The accounts were based on incidents that could happen to an undergraduate, such as visiting a doctor, studying for an exam, and interviewing for a job (see Appendix A for an example). For each story, both perceptual and emotional details were generated. The number of additional details of each type ranged from one to four. For each story, the number of additional "idea units" created was equivalent; that is, if two emotional detail units were created, two perceptual detail units were created as well. We were interested in people's believability ratings as a function of adding perceptual information and as a function of adding emotional detail. We therefore orthogonally crossed type of detail (perceptual, emotional) with level of detail (none added, added) to create four types of stories: no detail added, only perceptual details

added, only emotional details added, or both perceptual and emotional details added (hereafter referred to as both).

For any one participant, two accounts were presented in each form (i.e., no detail, perceptual detail, emotional detail, both). Across participants, each version of the story was assigned equally often to each condition. Likewise, each story appeared in each ordinal position equally often across participants in each condition. Complete counterbalancing of story with type of detail (i.e., story version) and ordinal position would be prohibitive. Hence, two sets of 64 story version/ordinal position combinations were created and half of each set was administered to the younger participants and half to the older participants. Each account was given a descriptive title (such as "Doctor's Office" or "Studying for the Exam") to remind participants of a particular account when they were required to rate its believability.

RESULTS AND DISCUSSION

We were interested in participants' believability ratings as a function of type of detail, age, and retention interval. To this end, we averaged each participants' ratings on the two stories containing each type of detail (i.e., no detail, perceptual detail, emotional detail, both—that is, both perceptual and emotional detail)¹ to obtain a single score for each type of added detail. We treated perceptual detail and emotional detail as within subjects variables with two levels each (absence, presence). The scores were submitted to a 2 (younger, older) \times 2 (immediate, delay) \times 2 (no perceptual detail, perceptual detail) \times 2 (no emotional detail, emotional detail) analysis of variance.

There was a reliable main effect of adding perceptual detail ($F [1, 124] = 3.84, MSe = 2.07$)², as well as a reliable main effect of adding emotional detail ($F [1, 124] = 35.97, MSe = 1.83$). There was no interaction. As expected, adding either type of detail to stories increased participants' tendency to believe these were accounts of actually experienced events. (See Figure 1.)

There was no significant interaction between age and type of detail. Contrary to expectations, the impact of emotional information was not greater for older than younger adults. This is somewhat surprising in

1. A materials error affected three of the younger participants' packets, in which incorrect versions of one of the eight critical stories were presented. When analyzing these participants' data, the three affected cells were filled with the mean score for the unrepresented condition (e.g., the mean score for younger participants in the delay condition for perceptual and emotional details).

2. An alpha level of .05 was used for all statistical analyses in this article.

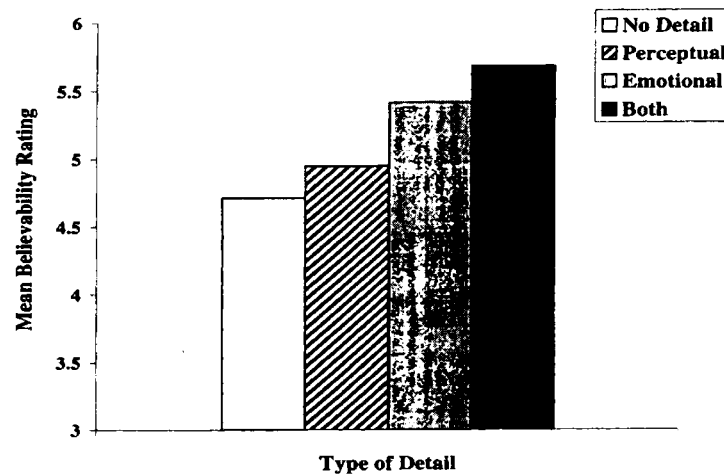


Figure 1. Mean believability ratings as a function of type of added detail (none, perceptual, emotional, or both perceptual and emotional), Experiment 1.

light of the previously discussed empirical findings regarding age differences in ratings of memory characteristics for perceived and imagined events (e.g., Hashtroudi et al., 1990).

A reliable main effect of retention interval ($F[1, 124] = 8.81, MSe = 2.36$) was superceded by a significant age by retention interval interaction ($F[1, 124] = 4.51, MSe = 2.36$). Planned comparisons showed that, while younger participants' ratings did not change over the retention interval (5.04 vs. 5.15 for immediate and delay, respectively), older participants' ratings increased when they were judging stories from memory—that is, after a brief delay (4.95 vs. 5.64 for immediate and delay, respectively). In other words, for older adults only, accounts were more likely to be judged as reports of actually-experienced events after a brief delay than when rated at the time of exposure. One possible reason for this interaction is that older adults may respond to their own perceived—or actual—poor memory of what others say by giving the benefit of the doubt, at least under conditions when they are predisposed to believe.

Finally, the delay by perceptual detail interaction was significant ($F[1, 124] = 3.78, MSe = 2.07$). This interaction came about because the immediate group's ratings of the accounts without perceptual information were lower than any other group.

One question that arises in interpreting the finding that increasing perceptual and emotional detail increases account believability is whether account length was a possible confound. Were our effects due

to type of detail added to the accounts, or could they reflect a belief that accounts of actually-experienced events tend to be longer? Danzer and Johnson (reported in Danzer, 1990) controlled for such a possibility in a conceptually similar experiment carried out with different materials. Their materials were designed so that half of the accounts in their longest versions (i.e., full detail) were still absolutely shorter than the remaining accounts in their shortest versions (i.e., no detail). The results showed that length did not account for differences obtained in believability ratings under the various detail conditions. Thus, it seems unlikely that length could account for our results in Experiment 1.

In summary, the key finding of Experiment 1 was that for both younger and older adults the perceptual and emotional content of accounts influenced their tendency to judge the accounts as those of actually experienced events. These judgments thus resembled personal reality monitoring with respect to the positive impact of perceptual and emotional information (e.g., Johnson et al., 1988).

EXPERIMENT 2

One critical issue is whether the obtained results generalize to all credibility assessments, or are instead particular to certain evaluative contexts. Philosopher H. P. Grice (1967/1989) described the default assumptions of everyday conversations. He suggested a cooperative principle, with a quality maxim that allows listeners to assume that speakers will not make statements unless they believe those statements to be true and have adequate evidence to support their belief. In effect, Grice was suggesting that we are predisposed to believe others in interpersonal reality monitoring (see also Gilbert, 1991). But, if conversation occurs in a situation in which the speaker is no longer expected to be obeying the cooperative principle, listeners may come to evaluate the speakers' assertions very differently (e.g., Bush, Johnson, & Seifert, 1994; Schwarz, Strack, Hilton, & Naderer, 1991; Wright & Wells, 1988). It seems reasonable that, if conversational expectations were shifted so that one is predisposed not to accept but to doubt a memory report, more emotional or perceptual details might make an account sound more contrived, causing one to doubt the report's truthfulness.

Findings generally consistent with this idea were obtained by Danzer and Johnson (reported in Danzer, 1990). Results suggested that the impact of perceptual and emotional information on account believability was context-sensitive. Participants (all younger adults) were required to rate the believability of simulated memory reports under three different sets of instructions. For all three conditions, perceptual and emotional content influenced ratings. However, the patterns of results for two of the conditions looked quite different from those of our

Experiment 1 in that adding detail to accounts failed to increase believability relative to the no detail version. Specifically, in one condition, accounts allegedly originated in transcripts of criminal trial testimony, where some of the speakers (those concealing guilt) would have a high motivation to lie. In another condition, participants were informed that the accounts came from honest students invited to the lab, but participants were explicitly cautioned that "people's memories are not always accurate, even when they are trying to be." An interesting similarity between these two sets of instructions is that both likely raise doubts about the accuracy of the accounts.

In Experiment 2, we examined the impact of suspicion on believability assessments. We used a modified version of the suspicion-inducing instructions employed by Danzer and Johnson (as reported in Danzer, 1990) and had participants (hereafter referred to as the high suspicion group) decide whether the accounts from Experiment 1 were essentially truthful or largely fabricated. We also included a condition to replicate the younger/immediate group from Experiment 1 (hereafter referred to as the low suspicion group). We were interested in how participants' judgments in each group varied as a function of type of detail added.

Also note that if, as anticipated, the added detail differentially affects believability ratings under low suspicion and high suspicion framing instructions in Experiment 2, we would have confirmation that judgments depend on qualitative characteristics of the information in combination with judgment context (e.g., how weights are assigned to qualities) rather than merely on report length.

METHOD

SUBJECTS AND DESIGN

Sixty-four younger adults who had not participated in Experiment 1 served as paid participants. They were Princeton undergraduate and graduate students, and undergraduate and high school students participating in summer programs at Princeton University.

Framing condition was a between subjects variable, with equal numbers of participants randomly assigned to the two framing groups (low suspicion instructions, high suspicion instructions). As in Experiment 1, type of detail (no detail, perceptual detail, emotional detail, both perceptual and emotional detail) was manipulated within subjects.

MATERIALS AND PROCEDURE

Except for the framing instructions, the materials and procedure were identical to those used for the immediate group in Experiment 1. The

accounts employed in Experiment 1 served again as simulated memory reports. A random half of the participants (i.e., low suspicion group) were given the "empathy experiment" cover story and rating instructions used previously. Recall that participants in this condition were told that all reports were of true events, the only question was whether or not the events were actually experienced by the teller (as opposed to their friend). The remaining participants (i.e., high suspicion group) were given a cover story designed to induce suspicion. They were told that the memory reports had been randomly excerpted from police interviews. These participants were led to believe that some of the accounts might have been fabricated by people with high motivation to lie (e.g., presenting a false alibi, protecting a loved one accused of a crime). They were asked to judge the believability of each account on an 8-point scale ranging from 1 ("largely fabricated") to 8 ("essentially truthful").

RESULTS AND DISCUSSION

We were interested in participants' believability³ ratings as a function of type of detail and framing instructions. Once again, the averages of the ratings of the two stories of each type (i.e., no detail, perceptual detail, emotional detail, both) were used in the analyses.

The mean believability ratings for Experiment 2 are shown in Figure 2. Because the two groups were making different kinds of believability ratings—that is, rating whether the account was of a personally experienced event (low suspicion) versus rating the veridicality of the account (high suspicion)—we conducted two separate ANOVAs. We examined the ratings in each framing condition individually, with perceptual detail and emotional detail as within subjects variables each having two levels. Separate analyses for each framing condition seemed especially appropriate given the fact that the two types of believability judgments

3. For ease of exposition, we use the word "believability" throughout the remainder of the paper to refer to both: people's tendency to believe that the information in the narrative was experienced by the person who told the story (under low suspicion instructions) and people's tendency to believe that the story told was a truthful account of an event (under high suspicion instructions). The two judgments (and, hence, the "truth" at issue) were somewhat different. For the low suspicion group the question centered on whether the narrative was a first-person account—that is, true to the speaker's experience—but the truth of the story (and the truthfulness of the teller), per se, was not in question. For the high suspicion group the question was directed at the truthfulness of the story (and thus the teller).

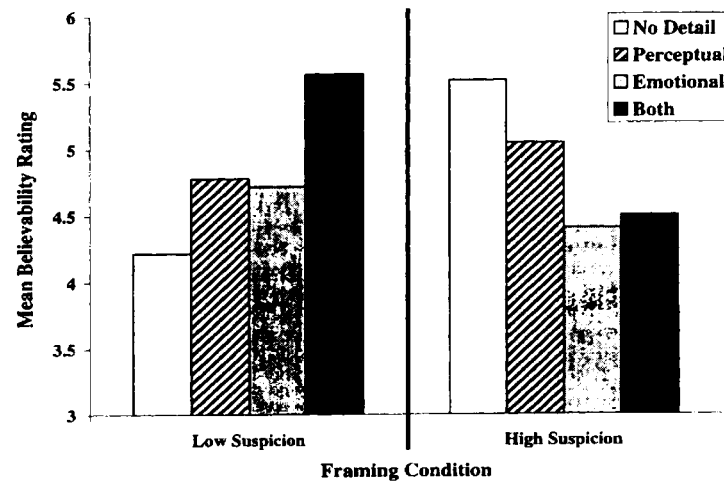


Figure 2. Mean believability ratings as a function of type of added detail (none, perceptual, emotional, or both perceptual and emotional) for each framing condition, Experiment 2.

clearly had different anchor points, as illustrated by the base rates for the no detail versions in each framing condition.

Analyses confirmed that in the low suspicion group, there was a significant effect of adding perceptual information ($F[1, 31] = 5.00$, $MSe = 3.17$) and a significant effect of adding emotional information ($F[1, 31] = 4.01$, $MSe = 3.28$). There was no interaction between the perceptual and emotional factors. Thus, as is evident from Figure 2, the general pattern of believability ratings under the low suspicion instructions resembled that found in Experiment 1. In both experiments, adding either type of detail to the base account increased believability, and the highest ratings were obtained in the condition combining perceptual and emotional details. Interestingly, whereas ratings of the low suspicion group increased with additional detail, those of the high suspicion group decreased. Analyses showed that in the high suspicion condition, adding emotional detail decreased believability ratings ($F[1, 31] = 10.40$, $MSe = 2.11$); and although adding perceptual detail decreased ratings somewhat as well, the effect was not significant, nor was the interaction between emotional detail and perceptual detail significant.

This pattern of results suggests that in making interpersonal reality monitoring judgments people weight various characteristics of memories differently, depending on their agendas or how the task is framed,

just as they do in evaluating the source of their own memories (e.g., Dodson & Johnson, 1993; Lindsay & Johnson, 1989; Marsh & Hicks, *in press*; Mather, Henkel, & Johnson, 1997).

The results are also consistent with the judgment and decision-making literature on framing effects (e.g., Shafir, 1993). Shafir found that in choosing between informationally enriched and impoverished options, people selected the enriched option more often than they selected the impoverished option when asked to choose one, and people rejected the very same enriched option more often than they rejected the impoverished option when asked to reject one. The results of Experiment 2 similarly represent a situation in which enriched detail information seemed to serve in the one case as a reason for belief, and in the other as a reason for disbelief.

EXPERIMENT 3

The results of Experiment 2 demonstrate the importance of judgment context to interpersonal reality monitoring. The very same information that on the one hand indicated to participants that a reported memory was personally experienced, on the other hand indicated that the memory was not experienced as described. Experiment 3 examined whether judgment context would influence believability judgments of elderly participants as well as younger adults. In addition, we were interested in more closely examining the extent to which the framing instructions influenced how participants made their believability ratings. Therefore, after participants had finished the objective believability-rating portion of the procedure, we simply asked them to tell us how they made their decisions (see Johnson et al., 1988).

According to the source monitoring framework (Johnson et al., 1993; Johnson & Raye, *in press*), the mix of heuristic and systematic processes used in reality monitoring depends on such things as judgment context. Heuristic processes assess qualities of the memory (e.g., perceptual detail) and systematic processes assess, for example, consistency and plausibility (see also Ross & MacDonald, 1997). Thus, one reasonable prediction is that participants in the low suspicion condition might rely more heavily (relative to the high suspicion group) on specific perceptual or emotional details, because these sorts of details generally are the input to the heuristic decision process in source monitoring under conditions where the default is to believe or to look for positive evidence. In contrast, participants in the high suspicion condition, who might be looking for reasons to disbelieve, might be expected to engage in more strategic or "reason/logic" based judgment processes (Johnson & Raye, 1981, *in press*).

METHOD

SUBJECTS AND DESIGN

see below for Experiment 1
Princeton undergraduates participated in exchange for course credit or a small monetary payment. In addition, ¹older adults from the same population as in Experiment 1 were paid for their participation. None of the participants had taken part in the earlier experiments.

The older adults were tested individually and the younger participants were tested in small groups ($n \leq 10$). As in Experiment 2, framing condition (low suspicion, high suspicion) was a between subjects factor, and equal numbers of younger and older participants were randomly assigned to each of the two experimental conditions. Type of detail (no detail, perceptual detail, emotional detail, both perceptual and emotional detail) was again manipulated within subjects.

MATERIALS AND PROCEDURE

The materials and procedure for the believability-rating task were identical to Experiment 2, with the obvious exception that there were both younger and older participants. In addition, after participants in this experiment had finished their believability ratings, they were asked to explain the basis for their decision. Specifically, they were given the following instructions (Note that the high suspicion group's instructions are presented in full with the changes for the low suspicion group in parentheses.):

We are interested in understanding the criteria people use in order to make judgments. In the space provided, please write down some of the things that influenced your decision. In other words, what swayed you to judge an account to be *essentially truthful (experienced by the speaker)* or *largely fabricated (not experienced by the speaker)*? Feel free to look back at your ratings, but, please do not change any of your answers. You may use examples from specific accounts to help explain your judgments.

RESULTS AND DISCUSSION

As illustrated in Figure 3, the main finding of Experiment 2 was essentially replicated; that is, the impact of perceptual and emotional details on the perceived truthfulness of an account depended on how the judgment was framed. To assess the nature of this impact, we once again conducted separate ANOVAs on the data from the low suspicion and the high suspicion groups, with age as a between subjects factor and

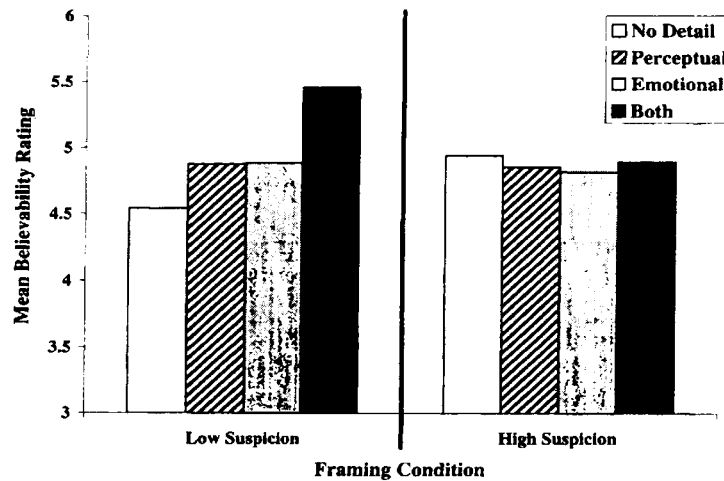


Figure 3. Mean believability ratings as a function of type of added detail (none, perceptual, emotional, or both perceptual and emotional) for each framing condition, Experiment 3.

perceptual detail and emotional detail as within subjects variables with two levels each.

Consistent with the previous experiments, analysis of the responses for the low suspicion condition revealed a reliable effect of adding perceptual detail ($F[1, 126] = 10.78, MSe = 2.46$) and also a reliable effect of adding emotional detail ($F[1, 126] = 12.92, MSe = 2.12$), and no significant interaction between perceptual and emotional detail.

In the high suspicion condition of Experiment 2 adding detail decreased believability ratings. In Experiment 3, there was no effect of adding detail in the high suspicion condition ($F_s < 1$ for both perceptual detail and emotional detail). There was a main effect of age in this condition such that the older participants gave overall lower believability ratings to the accounts ($F[1, 126] = 18.75, MSe = 3.39$)⁴. Age did not interact with type of detail ($F < 1$). Thus, it would appear that the older adults were generally more cautious in making their judgments in this particular high suspicion context.

4. *Ms* in the high suspicion condition, as a function of age, were: for the younger participants, 5.28, 5.26, 5.22, 5.17, for the no detail, perceptual detail, emotional detail, and both (perceptual plus emotional) detail versions, respectively; for the older adults, 4.61, 4.45, 4.42, 4.62, for the no detail, perceptual detail, emotional detail, and both (perceptual plus emotional) detail versions, respectively.

TABLE 1. Number of Participants' First Responses in Each Information Category as a Function of Framing Condition and Age (Exp. 3)

Information Category	Framing Condition			
	Low suspicion		High suspicion	
	Younger	Older	Younger	Older
Details	37	26	19	5
Feelings	14	5	1	1
Linguistic Features	7	10	5	4
Reasoning	2	13	32	43
Other	4	9	7	7

Note. $N = 251$. The "other" category includes miscellaneous comments that could not be easily categorized like "just guessed" or "I used my instincts."

Finally, we looked at participants' reports about their decisions. Although the majority of participants gave multiple reasons for their judgments, we limit the discussion here to their first response because not everyone cited more than one reason (cf. Johnson et al., 1988). We categorized participants' responses as generally centering on: details (participants said they focused on such things as: the presence/absence of perceptual details in general, the inclusion of a very specific detail such as the existence of a stained glass window, etc.); feelings (participants said they looked for the presence/absence of discussion of any discrete feelings or emotions, specific mention of a particular emotion like fear that would be consistent with the event, etc.); linguistic features (participants said they looked for such things as: unique linguistic cues like inclusion of the word "I," sentence complexity, etc.); reasoning (asking themselves how likely this event would be to occur in the real world, noting internal inconsistency in the story, asking themselves if the person had reason to lie, etc.); and other (i.e., miscellaneous comments that could not be easily categorized, e.g., "just guessed" or "I used my instincts").

Table 1 presents the number of participants whose first responses were assigned to each information category as a function of framing condition and age.⁵ The preponderance of participants' responses fell into two categories: details (34.7%) and reasoning (35.8%). A chi-square analysis confirmed that the pattern of responses was not independent of framing instructions, whether the analysis was conducted on the

5. Due to a clerical error, the narrative reports of four elderly participants were not collected (all in the high suspicion group). In addition, one elderly/low suspicion participant gave no justification, saying he could not verbalize the basis for his response. Thus, the discussion is based on the data of 251 participants (128 younger: 64 each in the low suspicion and high suspicion conditions; 123 elderly: 63 low suspicion and 60 high suspicion).

responses in all of the categories ($\chi^2(4) = 73.72$), or confined to only the detail and reasoning categories ($\chi^2(1) = 53.54$). Interestingly, as is readily apparent from Table 1, while participants in the low suspicion group tended to report focusing more on concrete details, such as perceptual features, than did the high suspicion group (49.6% of the low suspicion group vs. 19.3% of the high suspicion group), the high suspicion group was more likely than the low suspicion group to report engaging in reasoning in making their judgments (60.5% of the high suspicion group vs. 11.8% of the low suspicion group). Moreover, this basic pattern does not change if we consider participants' first three responses. Thus, it would appear that participants in the high suspicion group clearly were more suspicious, and that their suspiciousness influenced the information they considered when making judgments regarding the truthfulness of someone's reports. The fact that believability judgments in the high suspicion condition were uninfluenced by the manipulation of perceptual and emotional detail is consistent with these self-report data, suggesting these participants were relying more on systematic evaluation of consistency and plausibility than were participants in the low suspicion condition.

Johnson and colleagues (Johnson, 1988, 1991; Johnson et al, 1993; Johnson & Raye, 1981, 1998, in press) have suggested that, in evaluating their own memories, people tend to make heuristic judgments based on readily available information such as familiarity or the amount/quality of perceptual detail to make source attributions. When there is reason to doubt the veracity of one's heuristic judgment or when the decision is especially important (e.g., when the cost of an error would be great), people tend to engage in more strategic processes such as reasoning based on related prior knowledge, consistency or plausibility, or the extent to which related information comes to mind. The findings from Experiment 3 support the idea that similar factors influence attributions about other people's memories.

Interestingly, younger and older participants reported using details and reasoning in different proportions. Younger participants generally reported focusing more on concrete details such as perceptual features than did the older participants (43.7% of the younger participants vs. 25.2% of the elderly participants), and older participants were more likely than the younger participants to report engaging in reasoning (45.5% of older participants vs. 26.6% of younger participants). Nevertheless, details in both Experiments 1 and 3, had similar impact on believability ratings of younger and older adults. This raises an interesting question for future research: Are elderly less aware than younger adults of the impact details have on their judgments?

Another interesting observation was that the low suspicion group

was more likely to report using details regarding discrete and specific feelings and emotions (15%) than was the high suspicion group (1.6%). This finding for the low suspicion group is consistent with those of other reality monitoring studies showing that people use the presence of affect to infer that an event really happened (Johnson et al., 1988; Johnson & Suengas, 1989). Moreover, the present findings suggest that people may give greater weight to emotion when predisposed to believe than when predisposed to disbelieve. Finally, contrary to previous findings suggesting that the elderly tend to focus more on emotional details when making personal reality monitoring decisions (e.g., Hashtroudi et al., 1990), here younger participants were more likely to report using this information (11.7%) than were the elderly (4.9%). Nevertheless, the believability ratings (especially in the low suspicion condition) indicate that older adults were actually influenced as much as were younger participants by emotional detail.

GENERAL DISCUSSION

These experiments demonstrate the applicability of the source monitoring framework to understanding interpersonal reality monitoring. In particular, they show that the judgment context influences: (1) the weights assigned to various types of information, and (2) the mix of heuristic and systematic processes engaged in making source attributions (e.g., Johnson & Raye, 1981, 1998, in press). The very same information that on the one hand increased participants' confidence that a reported memory was truly experienced by the teller (Experiments 1, 2, and 3), on the other hand either reduced confidence that the event was experienced exactly as described (Experiment 2) or had no effect on such ratings (Experiment 3). In the low suspicion condition, adding perceptual and emotional detail to an account of an event increased participants' judgments that accounts were from people who actually experienced the events rather than from people reporting events experienced by a friend. In contrast, in the high suspicion condition, participants either reacted to added emotional or perceptual details as "too much of a good thing" (Experiment 2) or were not influenced by it at all (Experiment 3).

Clearly, the impact that perceptual and emotional details have on the attribution one makes about the source of an account depends on the context and the default assumptions surrounding the judgment. It is as if participants sometimes look for reasons to believe, and sometimes for reasons not to believe (cf. Shafir, 1993), or shift from more heuristic to more systematic evaluation processes (Johnson & Raye, 1981, in press; cf., Chaiken, 1987; Chaiken, Wood, & Eagly, 1996). The

self-report data from Experiment 3 were consistent with the idea of shifting from heuristic to systematic processing. Participants were more likely to report using detail in making believability judgments in the low suspicion condition and to report using reasoning in the high suspicion condition. Overall, these findings are consistent with the source monitoring framework, which posits that judgments regarding the origin of a memory are based on qualities such as the perceptual and emotional characteristics of the memories, and which emphasizes that characteristics are evaluated by heuristic and systematic processes in light of a person's metamemory assumptions, goals, and motives.

The patterns of believability ratings were generally the same for younger and older adults. However, one finding is particularly intriguing. Older adults' reports regarding the bases for their decisions indicated they were less likely to use amount of detail in making their decisions than were the younger adults. The objective believability ratings, on the other hand, suggest they were actually just as likely as younger adults to be influenced by the amount of detail. The possible mismatch between people's awareness of their source monitoring processes (e.g., what information is considered, what weights are assigned) and their actual attributions (cf., Nisbett & Wilson, 1977), potential age differences in the degree of this mismatch, and its practical ramifications, deserve further attention.

Other research on how people assess credibility has most often focused on situations closer to our high suspicion condition, such as jurors evaluating testimony. For example, work in this domain has considered the effects on credibility assessments of characteristics of the person reporting the memory (e.g., attractiveness), the speaker's style (e.g., perceived confidence), or motivation to deceive (e.g., Cutler, Penrod, & Dexter, 1990; Cutler, Penrod, & Stuve, 1988; Garcia & Griffitt, 1978; Leippe, Manion, & Romanczyk, 1992; Lindsay, Lim, Marando, & Cully, 1986; Riggio, Tucker, & Widaman, 1987; Wells & Leippe, 1981; Wells & Lindsay, 1983; Wells, Lindsay, & Ferguson, 1979). Our findings suggest there may be interactions between such variables and specific characteristics of memory reports (e.g., amount and type of detail).

Our results show that both perceptual and emotional detail incorporated into accounts of events have an impact on people's decisions regarding the origin of those accounts. However, the degree and even the direction of the impact depend on assumptions made about the intentions and motives held by the speakers of the accounts. It thus seems reasonable that interactions between the content of what is said and the context in which it is said could affect decisions made by jurors.

For example, details contained in an account given by a witness for the defense may affect a jury differently than the same details contained in an account given by a witness for the prosecution, depending on who jurors would like to believe (see also Bell & Loftus, 1988, 1989; but also see, Bell & Jones, 1994). Likewise, research has shown that jurors' attitudes toward the death penalty mediate their assessments of the quality of evidence delivered by prosecution and defense witnesses (e.g., Cowan, Thompson, & Ellsworth, 1984; Ellsworth, 1993; Thompson, Cowan, Ellsworth, & Harrington, 1984). More generally, there are likely to be other important unexplored interactions among variables affecting believability of memory reports (e.g., type of detail and confidence of the speaker; type of detail and cultural background of the person making the judgment).

In addition to informing the issue of how it is we go about assessing the veridicality of others' memories in our daily lives, or when we serve as jurors, the present results also suggest questions about interpersonal reality monitoring in applied domains in which "professionals" must evaluate the veracity of people's accounts of past events. For example, "statement reality analysis" (or "statement validity analysis") is a framework employed in German and Swedish courtrooms to expertly evaluate witness credibility. This approach exemplifies basic tenets of the source monitoring framework in that individuals engaged in statement reality analysis weigh characteristics of the information reported, such as the quantity and vividness of emotional and perceptual detail, taken together with their knowledge and beliefs about situational and motivational factors (e.g., Steller, 1989; Undeutsch, 1982, 1989).

Sporer (1997) recently compared forensic assessment of the contents of witnesses' reports, known as "criterion-based content analysis" (which often constitutes one part of a statement reality analysis), and reality monitoring criteria (e.g., Johnson & Raye, 1981). The results of this study showed that training with reality monitoring criteria allowed participant-judges to rate the veracity of "witnesses'" accounts at greater than chance levels, and slightly better than with the criterion-based content analysis criteria (although the procedure did not permit an unconfounded comparison of the two sets of criteria, see Sporer, 1997, p. 390). In fact, average differences between memories for actual and imagined events obtained under controlled laboratory conditions (e.g., Hashtroudi et al., 1990; Sporer, 1997; Suengas & Johnson, 1988) provides empirical support for the underlying assumptions of the statement reality analysis/criterion-based content analysis approach. Nevertheless, and perhaps more important, empirical work

on source monitoring also suggests caution in drawing firm conclusions about any particular memory because it is, of course, the atypical memories (e.g., imagined events high in perceptual detail) that are likely to be misattributed (Johnson et al., 1993; Johnson & Raye, 1998; Mitchell & Johnson, *in press*; see also Keogh & Markham, 1998).

Finally, the issue of alleged recovered repressed memories of childhood sexual abuse highlights questions regarding interpersonal reality monitoring in clinical contexts. The recent wave of court cases suggests that accurate assessment of the veracity of these accounts, not only by clients themselves, but by therapists, families, judges, and juries is an important issue. Given what we know about source monitoring more generally, it seems likely that people involved in these cases may sometimes make this assessment with overconfidence—that is, without sufficient understanding of either intra-individual or interpersonal reality monitoring processes (e.g., the impact of imagination, lax source monitoring criteria; e.g., Lindsay & Read, 1994, 1995). Empirical work is just beginning to investigate the accuracy of professionals' ability to monitor the veracity of other people's memories of past events (e.g., Landry & Brigham, 1992; Sporer, 1997; see also, Ceci, Crotteau Huffman, Smith, & Loftus, 1994; Ceci, Loftus, Leichtman, & Bruck, 1994). However, the present results highlight how important it is for therapists (as well as their clients) to fully appreciate how their own preconceptions might affect the weight they give to particular features of a client's account (e.g., affective content), and to its consistency and plausibility. Compared to someone with a predisposition to disbelieve, someone with a predisposition to believe is likely to assign a different set of weights to the same qualities of reported memories, and to engage a different distribution of heuristic and systematic processes.

There is much yet to be learned about how and in what circumstances laypeople and "expert" judges rely on various characteristics of reported memories, and how these qualitative characteristics are weighted differently depending on other variables. The issues of how credibility judgments are made and whether such judgments are accurate are related but dissociable. The informational characteristics of memory reports, the contexts or frames brought to bear, and the heuristic and systematic processes engaged, are relevant to both issues.

Appendix A

SAMPLE ACCOUNT

Note. Text common to all account versions is listed below. Brackets mark the locations in the text where additional details were inserted in the detail-enriched versions of the account (P = perceptual detail; E = emotional detail). Specific additional details are listed following the text of the no detail version. Additional details were either all ordinary or all unusual for any one participant; accounts were counterbalanced across participants on this factor.

DOCTOR'S OFFICE

No added detail

I walked into the waiting room [P1] [E1] when I gave my name to the nurse [P2] and sat down. [E2] [E3]. Then another nurse called my name [P3]. She led me to a room and told me that the doctor would be there in just a minute. [P4] Finally, the doctor come in and said that my tests had come out negative. [E4]

Perceptual Details (included in the perceptual detail version and the both [perceptual plus emotional detail] version):

- P1, ordinary:filled with people sitting around reading magazines.
- P1, unusual:filled with sports magazines and posters of racecars.
- P2, ordinary:who was sitting behind the counter
- P2, unusual:who was watering the plants
- P3, ordinary:in a loud voice.
- P3, unusual:in a thick Southern drawl.
- P4, ordinary:I sat on a table covered with a white paper sheet.
- P4, unusual:The phone rang several times while I was waiting in there.

Emotional Details (included in emotional detail version and both [perceptual plus emotional detail] version):

- E1, ordinary:I was kind of on edge
- E1, unusual:I was desperately worried
- E2, ordinary:I was getting more and more anxious,
- E2, unusual:I grew so anxious I could hardly breathe,
- E3, ordinary:but I tried not to panic.
- E3, unusual:and I felt like getting out of there fast.
- E4, ordinary:I felt so relieved!
- E4, unusual:I wondered if she was really telling me the whole truth.

REFERENCES

- Bell, B. E., & Jones, J. B. (1994). Providing minor details and the perception of honesty: Questioning the generality of trivial persuasion. *Psychology: A Journal of Human Behavior*, 31 (3/4), 26-29.
- Bell, B. E., & Loftus, E. F. (1988). Degree of detail of eyewitness testimony and mock juror judgments. *Journal of Applied Social Psychology*, 18, 1171-1192.
- Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, 56, 669-679.
- Bush, J. G., Johnson, H. M., & Seifert, C. M. (1994, August). The implications of corrections: Then why did you mention it? *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, Atlanta, GA.
- Ceci, S. J., Crotteau Huffman, M. L., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event. *Consciousness and Cognition*, 3, 388-407.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among pre-schoolers. *The International Journal of Clinical and Experimental Hypnosis*, 42, 304-320.
- Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3-39). Hillsdale, NJ: Erlbaum.
- Chaiken, S., Wood, W., & Eagly, A. H. (1996). Principles of persuasion. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 702-742). New York: Guilford.
- Conway, M. A., & Dewhurst, S. A. (1995). Remembering, familiarity, and source monitoring. *The Quarterly Journal of Experimental Psychology*, 48A, 125-140.
- Cowan, C. L., Thompson, W. C., & Ellsworth, P. C. (1984). The effects of death qualification on jurors' predisposition to convict and on the quality of deliberation. *Law and Human Behavior*, 8, 53-79.
- Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, 14, 185-191.
- Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior*, 12, 41-55.
- Danzer, A. (1990). *Judging the origin of other people's memories*. Unpublished manuscript, Princeton University, New Jersey.
- Dodson, C.S., & Johnson, M.K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, 106, 541-557.
- Durso, F. T., & Johnson, M. K. (1980). The effects of orienting task on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning and Verbal Behavior*, 19, 416-429.
- Ellsworth, P. C. (1993). Some steps between attitudes and verdicts. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 42-64). New York: Cambridge University Press.
- Finke, R. A., Johnson, M. K., & Shyi, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition*, 16, 133-137.
- Garcia, L. T., & Griffitt, W. (1978). Impact of testimonial evidence as a function of witness characteristics. *Bulletin of the Psychonomic Society*, 11, 37-40.
- Gilbert, D.T. (1991). How mental systems believe. *American Psychologist*, 46, 107-118.
- Grice, H. P. (1967/1989). Logic and conversation. In H. P. Grice (1989), *Studies in the way of words* (pp. 22-40). Cambridge, MA: Harvard University Press.
- Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging*, 5, 119-126.

- Johnson, M. K. (1983). A multiple-entry, modular memory system. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 17, pp. 81–123). New York: Academic Press.
- Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Oltmanns, & B. A. Maher (Eds.), *Delusional beliefs* (pp. 34–65). New York: Wiley.
- Johnson, M. K. (1991). Reality monitoring: Evidence from confabulation in organic brain disease patients. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury: Clinical and theoretical issues* (pp. 176–197). New York: Oxford.
- Johnson, M. K. (1997). Identifying the origin of mental experience. In M.S. Myslobodsky (Ed.), *The mythomaniacs: The nature of deception and self-deception* (pp. 133–180). Mahwah, NJ: Erlbaum.
- Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, 117, 371–376.
- Johnson, M. K., Hashtroudi, S. H., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28.
- Johnson, M. K., & Multhaup, K. S. (1992). Emotion and MEM. In S.-A. Christianson (Ed.), *The handbook of emotion and memory: Current research and theory* (pp. 33–66). Hillsdale, NJ: Erlbaum.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67–85.
- Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Sciences*, 2, 137–145.
- Johnson, M. K., & Raye, C. L. (in press). Cognitive and brain mechanisms of false memories and beliefs. In D. L. Schacter & E. Scarry (Eds.), *Memory and belief*.
- Johnson, M. K., & Suengas, A. G. (1989). Reality monitoring judgments of other people's memories. *Bulletin of the Psychonomic Society*, 27, 107–110.
- Keogh, L., & Markham, R. (1998). Judgements of other people's memory reports: Differences in reports as a function of imagery vividness. *Applied Cognitive Psychology*, 12, 159–171.
- Landry, K. L., & Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability of detecting deception in adults. *Law and Human Behavior*, 16, 663–676.
- Leippe, M. R., & Manion, A. P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology*, 63, 181–197.
- Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, 17, 349–358.
- Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology*, 8, 281–338.
- Lindsay, D. S., & Read, J. D. (1995). "Memory work" and recovered memories of childhood sexual abuse: Scientific evidence and public, professional, and personal issues. *Psychology, Public Policy, and Law*, 1, 846–908.
- Lindsay, R. C. L., Lim, R., Marando, L., & Cully, D. (1986). Mock-juror evaluations of eyewitness testimony: A test of metamemory hypotheses. *Journal of Applied Social Psychology*, 16, 447–459.
- Marsh, R. L., & Hicks, J. L. (in press). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, 25, 826–837.
- Mitchell, K. J., & Johnson, M. K. (in press). Source monitoring: Attributing (and misattributing) memories to sources. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory*. New York: Oxford University Press.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.

- Riggio, R. E., Tucker, J., & Widaman, K. F. (1987). Verbal and nonverbal cues as mediators of deception ability. *Journal of Nonverbal Behavior*, 11, 126-145.
- Ross, M. (1997). Validating memories. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd, (Eds.), *Memory for everyday and emotional events* (pp. 49-81). Mahwah, NJ: Erlbaum.
- Ross, M., & MacDonald, T. K. (1997). How can we be sure? Using truth criteria to validate memories. In M. S. Myslobodsky (Ed.), *The mythomanias: The nature of deception and self-deception* (pp. 181-201). Mahwah, NJ: Erlbaum.
- Schooler, J. W., Clark, C. A., & Loftus, E. F. (1988). Knowing when memory is real. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Volume 1: Memory in everyday life* (pp. 83-88). New York: Wiley.
- Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 171-181.
- Schwarz, N., Strack, F., Hilton, D. J., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "irrelevant" information. *Social Cognition*, 9, 67-84.
- Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, 21, 546-556.
- Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373-397.
- Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135-154). Dordrecht, The Netherlands: Kluwer.
- Stern, L. B., & Dunning, D. (1994). Distinguishing accurate from inaccurate eyewitness identifications: A reality monitoring approach. In D. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 273-299). New York: Springer.
- Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General*, 117, 377-389.
- Thompson, W. C., Cowan, C. L., Ellsworth, P. C., & Harrington, J. C. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior*, 8, 95-113.
- Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27-56). Deventer, The Netherlands: Kluwer.
- Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101-119). Dordrecht, The Netherlands: Kluwer.
- Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology*, 66, 682-687.
- Wells, G. L., & Lindsay, R. C. L. (1983). How do people infer the accuracy of eyewitness memory? Studies of performance and a metamemory analysis. In S. M. A. Lloyd-Bostock, & B. R. Clifford (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives* (pp. 41-55). New York: Wiley.
- Wells, G. L., Lindsay, R. C. L., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology*, 64, 440-448.
- Wright, E. F., & Wells, G. L. (1988). Is the attitude-attribution paradigm suitable for investigating the dispositional bias? *Personality and Social Psychology Bulletin*, 14, 183-190.
- Zaragoza, M. S., Lane, S. M., Ackil, J. K., & Chambers, K. L. (1997). Confusing real and suggested memories: Source monitoring and eyewitness suggestibility. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 401-425). Mahwah, NJ: Erlbaum.