Journal of Experimental Psychology:
Learning, Memory, and Cognition
1996, Vol. 22, No. 5, 1101–1122

0278-7393/96/$3.00

Memory for Faces: Evidence of Retrieval-Based Impairment

Paul D. Windschitl
Iowa State University

Four experiments investigated whether and how interpolated faces cause impairment to memories for related target faces. Participants viewed target faces and then saw a presentation of interpolated faces that were related to some of the targets. Modified tests, which offered target and novel faces as recognition alternatives, detected impairment effects after short retention intervals but not after 48-hr intervals, indicating that spontaneous recovery had occurred. For the interpolated presentations, some participants were misled to believe that the faces were the same as the targets, and others were informed that they were similar but different. The impairment and recovery effects were not moderated by participants' beliefs about the interpolated faces. The recovery effects suggest that interpolated faces affected the retrieval but not the storage of memories for targets, even for participants who were successfully misled about the interpolated faces.

Suppose you personally witnessed a stranger named Clarence perform a heroic act. The next day you see a newspaper account of the event in which a facial photograph of someone other than Clarence is inadvertently substituted for Clarence's photograph along with a caption indicating that this person was the hero. Would this newspaper photograph impair your subsequent ability to recognize Clarence's face? If so, how?

Important for any theory of memory is the issue of how new information affects memory for old information to which it is related. Two general research approaches have been taken in addressing this issue—the traditional retroactive-interference approach that focuses on the effects of interpolated learning on memory for target stimuli and a more recently formulated eyewitness-memory approach that focuses on how misleading postevent information influences witness memory reports. Studies within both of these areas have been interpreted as showing that new information can impair memory for old information, but questions remain about the specific nature of memory impairment. The present work addresses two questions related to memory impairment phenomena. The first concerns how new information might work to impair memory for old information. The second concerns whether memory for a human face, such as that of Clarence, is susceptible to impairment caused by a related face.

This article is based in part on a master's thesis submitted by the author to Iowa State University. A portion of this work was presented at the Fifth Annual Convention of the American Psychological Society, June 1993, in Chicago, Illinois. I am grateful to Gary Wells for his input on various aspects of this research, as well as to Carla Chandler and Robert Belli for their comments on drafts of this article. I am also grateful to Angelisa Braaksma, Susannah Carr, Kathy Fetterer, Mark Hankins, Dana Hartwigsen, Angie Kersting, Jeremy Kurth, Liberty Matt, Diane Peterson, Lynda Sanders, Sara Thier, Maurizio Visani, Valerie Wade and Dawn Marie Wurtele for their help with preparing materials and conducting experimental sessions.

Correspondence concerning this article should be addressed to Paul D. Windschitl, Department of Psychology, Iowa State University, Ames, Iowa 50011. Electronic mail may be sent via Internet to s1.pdw@isumvs.iastate.edu.

How Does New Information Impair Memory for Old Information?

Since the earliest retroactive-interference studies, researchers have offered numerous hypotheses to explain how new information might impair memory (see, e.g., McGeoch, 1942; Melton & Irwin, 1940; Postman, 1963). The proliferation of hypotheses continues, and such hypotheses have shaped a lively and long-running debate concerning the nature of memory impairment (Anderson & Spellman, 1995; Bekerian & Bowers, 1983; Chandler, 1991; Christiaansen & Ochalek, 1983; Loftus & Loftus, 1980). The impairment hypotheses can be classified into two groups—a storage-based group and a retrieval-based group. According to storage-based accounts of memory impairment, memory for old information can be rendered unavailable when postevent information alters or otherwise affects the stored memory traces for old information (Loftus, 1975, 1981a). The most radical storage-based account of impairment suggests that a memory trace for old information can be replaced by new information (Loftus & Loftus, 1980). Alternatively, retrieval-based accounts assume that impairment is a problem of inaccessibility. According to many of these accounts, intact memories for old and new information coexist in storage but interfere with each other at the time of retrieval (Bekerian & Bowers, 1983; Chandler, 1989, 1991; Christiaansen & Ochalek, 1983). For example, a response competition hypothesis suggests that memories for new and old information might compete for access to conscious recall (McGeoch, 1942).

Research concerning the nature of memory impairment phenomena has not provided a clear picture of how new information impairs memory for old information. Many researchers in the retroactive-interference tradition have tended to favor retrieval-based explanations, such as those involving response competition. Findings from recent retroactive-interference studies seem to be best suited for retrieval-based accounts of impairment (see Chandler, 1991; Wheeler, 1995). For example, in a retroactive-interference experiment, Chandler (1991) found that memories for target nature scenes were impaired when participants subsequently viewed interpolated nature scenes that were similar to the target scenes. The impairment was temporary, however. Spontaneous recovery was exhibited by the end of a 48-hr retention interval. As pointed out by Chandler, storage-based accounts of retroactive interference cannot explain such recovery from memory impairment given the assumption that altered or erased memory traces cannot be unaltered or unerased. In discussions of findings from eyewitness-memory research, however, Loftus has consistently favored storage-based interpretations to explain how postevent information, such as a mention of a yield sign, can affect participants' accuracy in reporting an actually viewed event detail, such as a stop sign (see Loftus, 1991; Loftus & Loftus, 1980; Loftus, Miller, & Burns, 1978; for opposing views see Bekerian & Bowers, 1983; Christiaansen & Ochalek, 1983). Also, Metcalfe (1990) has utilized a storage-based assumption of impairment in successfully modeling results from several postevent information studies with a distributed memory model called the composite holographic association recall model (CHARM). The storage-based assumption used by CHARM can be described as involving a composite process in which new information and old information are blended in one memory representation.

In short, findings from recent retroactive-interference studies have favored retrieval-based interpretations of memory impairment, and this stands in contrast to the storage-based interpretations often offered for postevent misinformation effects. At first glance this may seem problematic, but only if one assumes that there is only one way (either storage-based or retrieval-based) for new information to impair memory for old information. There may, however, be one or more critical differences between the retroactive-interference studies and postevent information studies, such that conditions in retroactive-interference experiments foster retrieval-based impairment, whereas conditions in postevent information studies foster storage-based impairment. One critical difference may involve the beliefs that participants hold about the relationship between the new (interpolated/postevent) information and the old (target/event) information. In the typical retroactive-interference study, participants are aware that the new information is different from the old information, whereas in postevent information studies, participants are misled to assume that the new information is consistent with the old information. The present work was motivated by the hypothesis that participants' beliefs about the relationship between new and old information might be a critical determinant of how the new information affects the memory for old information. Specifically, the present experiments tested the hypothesis that if new information is believed to be redundant with or the same as old information, the resultant memory impairment will be of a storage-based variety, whereas if the new information is believed to be different from old information, the resultant impairment will be of a retrieval-based variety.

This hypothesis is similar to, but distinct from, the ideas of Hall, Loftus, and Tousignant (1984), who argued that the detection or nondetection of a discrepancy between new and old information would affect whether memory impairment occurs (see also Tousignant, Hall, & Loftus, 1986). These authors reasoned that when a person does not detect discrepancies between the postevent and event information, the postevent information can be integrated into memory for the event information, resulting in impairment for the event memory. When, however, a person detects discrepancies between the postevent and event information, the person encodes the postevent information in a distinct way such that it can be easily discriminated from original event information, resulting in no memory impairment. For the present experiments it was assumed that discrepancy detection might be an important factor when new information is presented after old information, but in a slightly different way than what was proposed by Hall et al. It was thought that discrepancy detection might affect the type of impairment that results, rather than determining whether any impairment occurs. Although discrepancy detection might cause participants to believe that new and old information is different, this would lead to retrieval-based impairment rather than to no impairment at all. No discrepancy detection, however, might help participants to maintain a belief that the new and old information are the same, and this would lead to storage-based impairment.

Why might the perceived relationship between new and old information affect what type of memory impairment results? Assuming that there are no functional limits to the amount of information that can be stored in long-term memory, there seems to be no reason for a memory system to not give a new stimulus its own trace location unless the new stimulus is believed to be a second presentation of an already-stored memory. When a new stimulus is believed to be a second presentation of an already-stored memory, however, it seems plausible that the memory system would modify the original trace to accommodate the second presentation and store both presentations as instances of the same stimulus.

According to this logic, when a person encounters new information believed to be distinct from old information, the new information is likely to be stored in traces separate from the traces for the old information. When new and old information are stored in separate traces, storage-based forms of impairment, such as trace alteration or blending, would not occur. Retrieval-based forms of impairment, however, might occur under such conditions, particularly if the new and old information located in the coexisting memory traces were related (e.g., contained similar information). For storage-based impairment to occur, it seems likely that a person must believe that the new information is redundant or congruent with previously encountered information. Only then does it seem plausible that memory traces for old information would be modified to accommodate new information, thus resulting in storage-based impairment.

Reconsider the opening scenario and the question asking how your memory for Clarence would be impaired (if at all) by the inaccurate newspaper photograph. According to the present hypothesis, the answer to the question depends on whether you assumed the person in the photograph was Clarence. If you believed that the person was Clarence, your memory for Clarence might be altered to accommodate what you saw in the photograph. If, however, you thought the person in the photograph was not Clarence (perhaps you noticed certain discrepancies), you would not store what you saw in the photograph as part of the event you witnessed. This would preclude alteration of your memory for Clarence but leave your memory susceptible to retrieval-based forms of impairment.

Testing for Storage-Based and Retrieval-Based Impairment

There are two prerequisites to testing any hypothesis about when storage versus retrieval forms of impairment occur. First, memory impairment must be demonstrated. Many recent studies have used the modified test, an adaptation of the standard recognition test, to inspect for memory impairment (e.g., Belli, Windschitl, McCarthy, & Winfrey, 1992; Ceci, Ross, & Toglia, 1987; Chandler, 1989, 1991; McCloskey & Zaragoza, 1985; Zaragoza, 1987, 1991). On the standard test, both the event and postevent information are offered as response alternatives. Because the standard test includes postevent information as a response alternative, misinformation effects (i.e., effects in which misled participants are less accurate concerning critical event details) that are detected with the standard test could be attributable to factors other than memory impairment (McCloskey & Zaragoza, 1985). These factors could cause participants to change their test responses in reaction to postevent information, even when the new information had no effect on their memories for the event. For example, according to a response-bias interpretation, some participants may not have encoded the critical event information. Of such participants, those who were exposed to postevent information would be more likely to select the postevent information on the test than would those who were not exposed to postevent information. According to a social-demand interpretation, participants who remember both event and postevent information might have a tendency to select postevent information on the standard test because of perceived social demand. To control for interpretations involving response bias, perceived social demand, and other factors, the modified test excludes postevent information as a response alternative and offers event and novel information as the only possible responses. Because of its ability to control for alternative interpretations, the modified test has become an accepted test for detecting memory impairment (see Belli et al., 1992; Chandler, 1989, 1991, for recent discussions).

The second prerequisite for testing a hypothesis about when storage-based versus retrieval-based impairment occurs is that a method for distinguishing between the two forms of impairment must be used. One method that can partially fulfill this prerequisite involves measuring for memory impairment at multiple time periods. If spontaneous recovery from memory impairment is observed, retrieval-based hypotheses are implicated. As mentioned above, storage-based hypotheses, as they are currently specified, cannot account for spontaneous recovery (see Chandler, 1991). Storage-based hypotheses, therefore, can be unambiguously ruled out as explanations for impairment that decreases over time. Alternatively, if memory impairment persists over multiple time periods, storage-based hypotheses become the most plausible candidates for explaining the impairment. Unfortunately, this method remains imperfect; the observed persistence of memory impairment can never lead one to fully rule out retrieval-based accounts of the impairment (see Loftus, 1981b, for discussion). As explained shortly, however, the design of the present experiments offered a unique opportunity for storage-based impairment to be identified as such.

The present experiments used the modified test at multiple time periods to investigate the hypothesis that a person's belief about the relation between old and new information affects how his or her memory for the old information is impaired. Participants viewed target faces and then viewed interpolated faces that were similar in appearance to many of the target faces. Each participant was led to believe either that the interpolated faces were the same as or that they were different from the target faces. It was expected that participants' memories for the target faces would be impaired by the interpolated faces, but that the form of the impairment would be dependent on whether they were told that the interpolated faces were the same or different from the target faces.

In accord with the hypothesis outlined above, participants in the told-different condition were expected to experience retrieval-based impairment. It was also thought that this expectation would be confirmed by the observation of spontaneous recovery (similar to a result observed by Chandler, 1991). Specifically, it was thought that told-different participants would exhibit memory impairment after a short retention interval but that the impairment would dissipate before the end of a 48-hr retention interval. Alternatively, participants in the told-same condition were expected to experience storage-based impairment. Thus, for these participants, memory impairment was expected to remain stable over time.

As stated above, observing persistence in impairment over multiple time periods cannot ensure that the impairment was storage based. The design of the present experiments, however, allowed for a pattern of results in which impairment suffered by told-same participants not only could persist across multiple time periods but also could persist over a time period during which told-different participants experienced recovery. This type of finding would suggest that the memory impairment observed for the two groups of participants was qualitatively distinct, such that the impairment mechanisms affecting told-same participants allowed a release from impairment after a given retention interval, but the impairment mechanisms affecting told-different participants precluded a release from impairment at that time period. If the impairment suffered by told-same participants was found to persist over a substantial retention interval, the most plausible (albeit not the exclusive) explanation for their impairment would be to assume that the impairment was storage based.

Faces as Stimuli

Human faces were used as the stimuli in the present experiments for two reasons. One reason was that faces are conducive to the task of manipulating people's beliefs about whether old and new stimuli are the same or different. Faces are highly complex stimuli that cannot easily be described. These qualities make it difficult for a person to make a decision about whether a newly presented face is the same or different from a similar face seen earlier. Thus, using faces as stimuli in the present experiments made it less likely that participants in the told-same condition would notice the subtle discrepancies between target and interpolated stimuli and conclude that the two stimuli were different.

The second, more significant reason for using faces as stimuli was to test whether memories for faces are susceptible to impairment. Some researchers have interpreted studies on prosopagnosia, face inversion, and people's impressive ability to recognize large numbers of faces as support for the position that face recognition involves a special processing system (e.g., Whiteley & Warrington, 1977; Yin, 1969, 1970). One derivation from this position is the idea that memories for faces might not be affected by impairment that plagues memories for other stimuli (see Davies & Christie, 1982; Davies, Shepherd, & Ellis, 1979).

A small number of memory-for-faces studies have shown that interpolated faces or misinformation can affect participants' performances at identifying target faces (e.g., Deffenbacher, Carr, & Leu, 1981; Jenkins & Davies, 1985; Laughery, Alexander, & Lane, 1971; Loftus & Greene, 1980). Like the effects found in many postevent information studies, however, the effects in these memory-for-faces studies can be attributed to factors other than memory impairment. In memory-for-faces studies involving misleading postevent information or misleading composites, the misleading information was included in alternatives on the memory tests. Although exposure to misleading information has been shown to affect test performances in such studies, these effects are attributable to misled participants trusting the postevent information over their own (unimpaired) memories for the target faces.

In the memory-for-faces studies concerning more traditional forms of retroactive interference, other factors, such as those involving source misattributions (see Lindsay & Johnson, 1989) can account for observed effects. Participants in Deffenbacher et al.'s (1981) study saw 21 target faces and then either judged 21 interpolated faces as targets or nontargets (experimental condition) or performed a digit-cancellation task (control condition). On a subsequent recognition test, participants saw a series of target, interpolated, and new faces and identified which faces were seen as targets. Relative to participants in the control condition, those in the experimental condition were less accurate on the test. Although retroactive interference was demonstrated in this study, the effect may have been the result of source misattributions rather than any memory impairment. Participants in the two conditions may have been equally able to recognize a face as seen or not seen, but only those in the experimental condition responded incorrectly because of confusion as to whether the face was seen in the target or interpolated presentation. Source misattributions can account for interference effects on tests that require participants to correctly determine when or where they had seen a given test face. In a recent review of studies involving retroactive interference and face memory, Deffenbacher (1991) suggested that confusion regarding the correct circumstances in which a face was encountered is the primary cause of retroactive interference effects observed in studies using recognition tests that include interpolated faces.

Criterion shifting is another factor that can account for the retroactive interference effects found in other memory-for-faces studies. In studies conducted by Laughery and his colleagues, participants' performances at identifying a previously seen target face suffered as the number of searched-through interpolated faces increased (Laughery et al., 1971; Laughery, Fessler, Lenorovitz, & Yoblick, 1974). Davies et al. (1979) argued that participants in those studies simply became more reluctant to make an identification as they searched through increasing numbers of interpolated faces. In their own study, Davies et al. (1979) found that when criterion shifting was controlled for, exposing participants to interpolated faces had no effect on performance at identifying previously seen targets.

To date, it appears that no memory-for-faces study has isolated the question of whether recognition of a face is impaired by exposure to a related interpolated face. Furthermore, in studies that controlled for certain alternative interpretations, such as the study by Davies et al. (1979) and some eyewitness identification studies, results suggest that interpolated faces might not affect memories for target faces (see Brigham & Cairns, 1988; Cutler, Penrod, & Martens, 1987a, 1987b). These studies, however, used between-subjects designs, some of which may have lacked a sufficient level of power to detect existing impairment. Also, the similarity between the target and interpolated faces used in these studies was not high. Given the importance of similarity in interference phenomena (McGeoch & McDonald, 1931), the above studies might not have provided conditions conducive to observing memory impairment. Thus, the question of whether memories for faces are susceptible to impairment is far from settled. The present experiments used the modified test to inspect for impairment and control for source misattributions, response biases, and other alternative interpretations. A partially within-subjects design was used to enhance detection of any existing impairment. Also, the interpolated faces used in the experiments were somewhat similar to related target faces. It was thought that these conditions would provide a reasonable test of whether memories for faces can be impaired by interpolated faces.

Experiment 1

Method

Overview

After rating 66 target faces on attractiveness, participants were given information about the upcoming presentation of new (interpolated) faces. Participants in the told-same group were misled to believe that the interpolated faces would be the same as the previously seen target faces, whereas those in the told-different group were correctly informed that faces would be different from, but similar in appearance to, the target faces. All participants then rated 33 interpolated faces on apparent honesty. Each interpolated face was matched, based on similarity, to one of the target faces seen previously. Thus, half of the target faces (the experimental-condition targets) were later followed by a matched face in the interpolated presentation, whereas the other half of the target faces (the control-condition targets) were not followed by a matched face in the interpolated presentation. A modified test (i.e., one that did not include interpolated faces) was given to participants either 10 min or 48 hr after the interpolated presentation.

Participants and Design

The participants were 96 Iowa State University students who volunteered to participate for extra credit in an introductory psychology class. The design was a 2 × 2 × 2 factorial with one within-subjects variable and two between-subjects variables. Each participant served in both the experimental and control conditions of the experiment. A given participant served in either the told-same or told-different group. Half of the participants in each instruction group were tested after a 10-min delay (between the interpolated and test phase) and half were tested after a 48-hr delay.

Materials

All of the target, interpolated, and test faces were presented on color video screens and were originally videotaped from still photographs found in various books and dated magazines. The faces were collected in sets of 3, such that the members of each set were somewhat similar in appearance and could be labeled with a common descriptor. For example, the "brides" set contained three similar-looking individuals (Brides A, B, and C), each wearing similar wedding attire. Along with 66 sets of faces (198 total faces), 20 single faces were also collected to be used as practice faces. Video presentations were assembled from this collection of faces.

The target presentations contained a total of 86 faces—the 20 practice faces followed by one target face from each of the 66 sets. The faces in the target presentation appeared as head-and-shoulder portraits that were surrounded by a black border. Below each portrait appeared a descriptive label.[1] For example, the label *Bride* accompanied the portrait of a woman wearing a wedding dress.

For the interpolated presentations, the black border was expanded such that only the face of a pictured individual was displayed. Head-and-shoulder views were not used in the interpolated and test presentation because such views would have supplied participants with nonfacial cues (e.g., type of wedding dress) that might have been used as a basis for responses at test or that might have led participants in the told-same condition to reject the idea that the target and interpolated faces were the same. Participants had to rely on the label and a few remaining cues (e.g., age and sex) to draw the connection between the target and interpolated faces. As in the target presentations, a descriptive label appeared below each face in the interpolated presentations. A presentation consisted of 10 practice faces, followed by 33 interpolated faces. Each interpolated face was similar to (and from the same set as) one of the target faces.

The test presentations contained 66 pairs of faces, 1 pair from each set. The target face from each set appeared next to the novel face from the set (i.e., the face that was not seen in the target or interpolated presentations). A black border separated and surrounded the faces, and the label common to the face set appeared below each pair. For example, in the bride set, if Bride A was displayed as a target and Bride B was displayed in the interpolated presentation, then the faces of Brides A and C were shown on the test presentation along with the label *Bride*. Figure 1 illustrates how faces appeared on the target, interpolated, and test presentations.

Randomization and counterbalancing schemes were used to determine the selection and ordering of faces used in target, interpolated, and test presentations. A counterbalancing scheme ensured that in all conditions of the experiment, the three faces from any given set served as targets equally often and as interpolated and novel faces equally often. Also, the face sets (e.g., brides) were used equally often in each of the eight conditions of the experiment. On the test tape, the locations of the target faces were randomly assigned with the restriction that the targets appeared on the left and right with the same frequency.

Procedure

Participants were tested in groups of two. Each participant was seated at his or her own desk on which rested a 33-cm color video monitor. Initial instructions indicated that the experiment involved rating numerous people on various dimensions. The target presentation was introduced, and participants were asked to judge the attractiveness of persons appearing on their video monitor. Participants placed a mark along an attractiveness continuum ranging from *very unattractive* to *very attractive* for each target face. Each target face was displayed for 8 s. During a 2-min delay between the presentation of targets and the introduction to the interpolated presentation, participants read cartoons. The introduction of the interpolated presentation differed depending on whether the participant was in the told-same or told-different group.

Told-same. Participants in the told-same group were misled to believe that they would see people from the target presentation again. The experimenter told the participants that they would again rate the people that they had just rated, this time using a different dimension. Additional instructions read by participants began with the following: "Now you will see half of these people again. Please rate each person again. . . ." The instructions then informed participants that they should indicate, on a continuum line anchored by *very dishonest* and *very honest,* how honest each of the following people appeared to be. The first 10 faces that appeared on the interpolated videos were included for the specific purpose of strengthening participants' beliefs that all of the faces that appeared in the interpolated presentation had already been seen in the target presentation. These first 10 faces did, in fact, appear within the first 20 portraits of the target presentation, and they were labeled as they were in the target presentation. In an additional attempt to strengthen participants' beliefs, the experimenter showed the participants a misleading illustration of what they were about to see. The experimenter held up two facial photographs; the first was said to be an example of what was seen in the target presentation, and the second was said to be an example of what would be seen in the interpolated presentation. The two example photographs were obviously of the same individual, pictured in a slightly different way with the latter example being more of a close-up of the individual's face.

Told-different. Participants in the told-different condition were correctly informed that the people they were about to see were different from the people in the target presentation. The experimenter told them that although the people in the interpolated presentation might look very similar to people in the target presentation, only the first 10 people were actually in the target presentation. Instructions read by the participants began with the following: "Now you will see a different set of people. These new people look very similar to the individuals in the first video and have the same labels as people in the first video. For people in this video. . . ." These instructions went on to describe the same honesty rating task that told-same participants read

[1] Given that this experiment was testing for what might be called item-specific impairment, steps were taken to optimize the conditions for interference between a given target face and a specified interpolated face. The faces were collected in sets such that the members would be associated and somewhat similar in appearance. The descriptive labels were included to strengthen the association participants made between a target face and an interpolated face. The linkage is also important for reducing the unmeasured interference among target faces, among interpolated faces, and among target and interpolated faces of different sets. Use of head-and-shoulder views in the target presentation enabled participants to use visual cues (e.g., a wedding dress) to establish a connection between the target face and the given label (e.g., *Bride*).

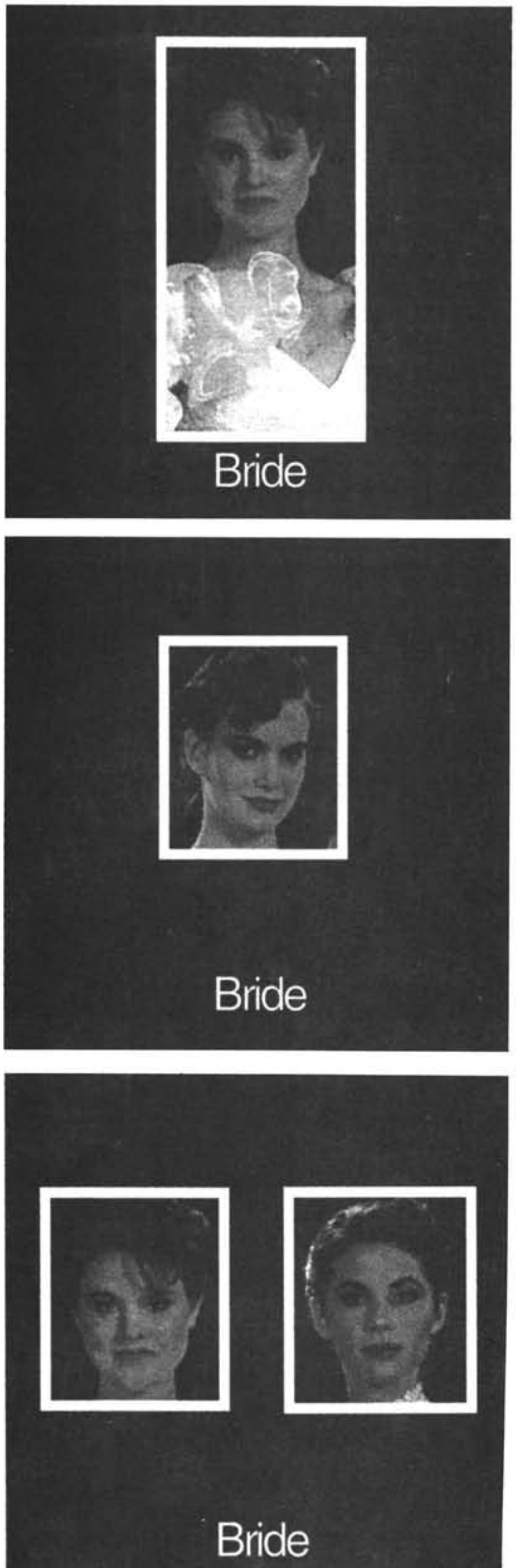

Figure 1. An example of how faces appeared in the target (top), interpolated (middle), and test presentations (bottom).

about. The final written instruction reminded participants that the first 10 faces of the interpolated videotape did appear in the target presentation. As in the told-same condition, the experimenter showed the participants an illustration. The photographs used in the illustration, however, were obviously portraying two different people.

After being given the told-same or told-different information, all participants viewed the interpolated presentation at a rate of 8 s per face. The presentation contained 10 faces already seen at the beginning of the target presentation, followed by 33 faces that were matched (i.e., were members of the same set and shared a common label) with 33 of the target faces. The 33 targets that were followed by a matched face in the interpolated presentation constituted the experimental-condition targets. The 33 targets that were not followed by a matched face constituted the control-condition targets.

After the interpolated presentation, participants in the 10-min group were immediately given a cartoon ranking task. They worked on that task for 10 min before the memory test was introduced. Participants in the 48-hr group were dismissed. When they returned 2 days later, they too worked on the cartoon ranking task for the 10 min prior to the memory test. In the introduction to the test, all participants were informed that the test would contain numerous pairs of faces with one label below them, and that they had seen one face from each pair. Participants were told to choose the face that appeared at any time during the experiment and, if unsure, to make a "best guess." Choices were indicated by circling either an *L* (left) or *R* (right) appearing next to each label on a test form. Each test pair was displayed for 10 s.

Results and Discussion

For each participant, a control-condition score (based on the number of control targets correctly recognized) and an experimental-condition score (based on the number of experimental targets correctly recognized) were computed. The mean control scores and experimental scores are displayed by groups in Table 1. Table 1 also contains memory impairment scores, which were calculated by subtracting the experimental scores from the control scores within each group. The control and experimental scores were submitted to a 2 (told-same/told-different) × 2 (delay) × 2 (control/experimental) analysis of variance (ANOVA) with the control/experimental variable as a within-subjects variable. Unless otherwise noted, a two-tailed alpha level of .05 was used for all of the analyses reported below.

Table 1
Mean Control-Condition and Experimental-Condition Performance (in Percentages) by Groups in Experiment 1

Group	Control condition		Experimental condition		Impairment effect (%)
	M	*SEM*	*M*	*SEM*	
10-min delay					
Told-same	78.8	1.4	74.4	1.8	4.4
Told-different	79.4	1.7	75.8	1.4	3.6
48-hr delay					
Told-same	69.1	1.9	69.7	2.1	−0.6
Told-different	72.1	1.9	73.4	2.1	−1.3

Note. The means are based on the percentages of the 33 targets correctly identified. The sample size for each mean is 24. The impairment-effect column indicates the mean differences between the control and experimental performances.

Overall, the participants correctly identified 74.8% of the control-condition targets (M = 24.7 of 33, SEM = 0.32) and 73.3% of the experimental condition targets (M = 24.2 of 33, SEM = 0.31). The ANOVA indicated that the main-effect difference between control scores and experimental scores was not significant, $F(1, 92) = 2.99$, $MSE = 4.02$, but this result must be interpreted in light of a significant Delay × Control/Experimental interaction, $F(1, 92) = 7.88$, $MSE = 4.02$. The interaction suggests that the differences between control scores and experimental scores were larger for the 10-min group than for the 48-hr group. Participants in the 10-min group correctly identified 79.1% of the control targets (M = 26.1, SEM = 0.36) but only 75.1% of the experimental targets (M = 24.8, SEM = 0.38). This difference in performance was significant, $F(1, 92) = 10.29$, $MSE = 4.02$, and it indicates that participants' memories for target faces were impaired when those target faces were followed by similar-looking interpolated faces. Although the impairment effect was robust in the 10-min group, no impairment was detected in the 48-hr group. For the 48-hr group, the experimental scores (M = 23.6, SEM = 0.49) were trivially higher than control scores (M = 23.3, SEM = 0.45), $F(1, 92) < 1$.[2]

The above results replicate Chandler's (1991) finding in which a memory-impairment effect was robust after 10 min but negligible after 48 hr. In Chandler's studies, however, no participants were led to believe that the interpolated stimuli were the same as the target stimuli. For the present experiment it was predicted that both the told-same and told-different participants would exhibit memory impairment at 10 min. However, it was thought that the impairment suffered by told-same participants would be storage-based and would persist through a 48-hr delay, whereas the impairment suffered by told-different participants would be retrieval-based and would dissipate over a 48-hr delay. Thus, it was thought that the three-way interaction would be significant with both the told-same and told-different participants exhibiting memory impairment in the 10-min group, but only the told-same participants exhibiting impairment in 48-hr group.

Contrary to this prediction, the three-way interaction was not significant, $F(1, 92) < 1$. Told-same and told-different participants exhibited similar levels of memory impairment within delay groups. As predicted for the 10-min group, both the told-same and told-different participants exhibited significant levels of memory impairment, as evidenced by higher control scores than experimental scores, $F(1, 92) = 5.99$, $MSE = 4.02$, for told-same participants, and $F(1, 92) = 4.36$, $MSE = 4.02$, for told-different participants. Unexpectedly, however, told-same participants in the 48-hr group did not exhibit memory impairment. For both the told-same and told-different participants tested at 48 hr, control scores and experimental scores did not differ, both $Fs(1, 92) < 1$. Analysis of the main effect for the told-same/told-different factor indicates that the overall test performance of told-same participants was not significantly different from the performance of told-different participants, $F(1, 92) = 1.86$, $MSE = 12.96$. Also, neither the Told-Same/Told-Different × Delay interaction, nor the Told-Same/Told-Different × Control/Experimental interaction were significant, both $Fs(1, 92) < 1$.

The results indicate that both the told-same and told-different participants exhibited significant levels of memory impairment when tested at 10 min but not at 48 hr. This pattern suggests that participants experienced a release from impairment sometime between 10 min and 48 hr after exposure to the interpolated faces. Although such a finding is consistent with the hypothesis that told-different participants would experience impairment of the retrieval-based variety, it is at odds with the hypothesis that told-same participants would exhibit impairment of the storage-based type. For both the told-same and told-different groups, the memory impairment detected after the 10-min interval must have been retrieval based. Storage-based hypotheses as they are currently defined cannot account for the recovery from impairment observed over the 48-hr delay.

In terms developed in the interference literature, the present pattern of results constitutes a *relative spontaneous recovery* (see Brown, 1976). From experiments within the paired-associate paradigm, numerous researchers reported finding that retroactive interference for A–B pairs caused by studying A–C or C–D pairs tended to lessen over time (e.g., Birnbaum, 1965; Postman, Stark, & Fraser, 1968; Underwood, 1948; see Wheeler, 1995, for a recent example). In those experiments, participants who had studied both A–B and then A–C or C–D word pairs (experimental condition) were less able to recall the A–B pairs after short retention intervals than were other participants who had studied only the A–B word pairs (control condition). After long retention intervals (e.g., 24 hr), however, the difference in A–B recall between the experimental and control groups dissipated. The term *absolute spontaneous recovery* was used to describe instances in which, between the short and long retention intervals, there was an improvement in A–B recall for experimental participants. The term relative spontaneous recovery was used to describe instances in which, between the short and long retention intervals, there was a decrease in A–B recall for experimental participants that was less than the decrease for the control participants (see Brown, 1976). In the present experiment, recognition performance for both the experimental and control items decreased from 10 min to 48 hr, but the decrease was significantly smaller for experimental items than for control items.

There are several possible explanations for how a memory suffering from impairment can be spontaneously recovered. Concerning the paired-associate paradigm, Underwood (1948) suggested that A–B associations are unlearned during A–C learning, but, like extinguished conditioned responses, recover strength with the passage of time. Postman et al. (1968) suggested a response-set suppression hypothesis in which a tendency to give the most recently learned response dissipates over time. More appropriate for the present results might be a competition hypothesis that suggests that a strong memory for new information can, at retrieval, compete with a coexisting memory for old information; but as time passes, the memory for the new information weakens and is less likely to compete for retrieval (see Chandler, 1991).

Although a demonstration of relative spontaneous recovery

[2] The power to detect a memory impairment effect in the 48-hr group was about .78 (assuming a one-tailed alpha of .05 and using the effect size estimate from the 10-min group, Cohen's d = .49).

suggests that there was a release from memory impairment, there is an alternative explanation for some relative recovery effects. The alternative explanation assumes that experimental performance at the short retention interval was near a floor level (or functional floor level) and that it was reduced to the floor level over the long retention interval. Hence, experimental performance would show little or no decline between the short and long retention intervals. Control performance, however, would not be near the functional floor level at the short retention interval, and, consequently, control performance would show a greater decline between the short and long retention intervals. Regarding the present experiment, one could argue that the interpolated faces reduced (through memory impairment) the experimental performance to a near-floor level on the 10-min test. During the 48-hr delay, extraexperimental forgetting caused both the control and experimental performance to fall to a floor level (or functional floor level). Although this forgetting caused a large decrease in control performance, it caused only a small decrease in experimental performance, because the experimental performance was already close to the floor level (see Koppenaal, 1963, for a discussion of negative accelerated unlearning).

This type of explanation for relative recovery effects is important to consider, because it does not assume the recovery of any impaired memories. It allows for the possibility that the impairment observed on the 10-min test still "exists" after 48 hr, but that no memory impairment is detected on the 48-hr test because both control and experimental performances have reached a functional floor level. Hence, storage-based hypotheses of impairment can be entertained as long as this explanation of a recovery effect is plausible.

There are two instances in which an explanation involving floor-performance levels can be ruled out as an account for a recovery effect. First, such an explanation cannot account for an absolute spontaneous recovery effect (i.e., an instance in which experimental performance is shown to increase over time). Second, explanations involving floor-performance levels can be ruled out by demonstrating that the functional floor level falls below the experimental-condition performance exhibited by participants at the long retention interval. The following two experiments were designed to disambiguate the recovery effects observed in Experiment 1 by testing for absolute recovery effects (Experiments 2 and 3) and by investigating the functional floor level for the relevant materials and design (Experiment 3).

Experiment 2

The most straightforward evidence for the dissipation of impairment is an absolute spontaneous recovery effect. As pointed out by Postman et al. (1968), absolute recovery effects will only be observed when the amount of recovery occurring in a given time period (e.g., between tests given at 10 min and 48 hr) significantly exceeds the amount of extraexperimental forgetting occurring in that time period. Experiment 2 contained two modifications that provide a better opportunity for recovery to exceed, or at least equal, the extraexperimental forgetting that occurs over a given time period.[3] First, to enhance the memory impairment effect, participants in Experiment 2 viewed the interpolated faces twice. The number of target faces viewed was reduced from 66 to 44 to compensate for the increase in the number of times interpolated faces were seen. Second, to avoid strong delay effects (caused by extraexperimental forgetting), a 45-min-delay group was included. It was thought that release from memory impairment might occur sometime between 10 and 45 min after seeing the interpolated faces and that extraexperimental forgetting over this interval would be minimal. A 48-hr group was also included, however, because it was considered possible that retrieval-based impairment might not subside until after the 45-min delay.

Experiment 2 also contained changes involving the told-same/told-different manipulation. The possibility was considered that the null effects involving the told-same/told-different variable were a result of a weak manipulation. Participants' postexperiment comments suggested that this was not the case; told-same participants were generally surprised to learn, during debriefing, that the interpolated faces were different from the target faces. Nevertheless, a small modification to the interpolated presentations was introduced to strengthen the manipulation, and a question was added to assess the adequacy of the manipulation.

Method

Participants and Design

The participants were another 144 students from the same pool used in Experiment 1. Except for the addition of the 45-min group, creating a 2 × 3 × 2 factorial, the design of the experiment resembled that of Experiment 1.

Materials and Procedure

The materials and procedures used in the experiment were identical to those used in Experiment 1, except for the following changes. The number of introductory (nontarget) faces presented at the beginning of the target presentations was reduced from 20 to 8. The number of introductory faces on the interpolated presentations was reduced from 10 to 5. The numbers of target and interpolated faces seen were reduced from 66 to 44, and 33 to 22, respectively.

To strengthen the told-same/told-different manipulation, the introductory faces on the interpolated presentations differed between the told-same and told-different groups. Told-different participants saw five faces that they had not seen before. Thus, the instructions used in Experiment 1 to warn them that the initial faces of the presentation had already been shown in the target presentation were omitted. Alternatively, the told-same participants saw five faces that had appeared at the beginning at the target presentation.

All participants viewed the interpolated faces twice. The procedures for viewing and rating the first presentation of the interpolated faces

[3] For Experiment 1, the upper limit on the amount of spontaneous recovery that could have occurred was somewhat small given that the impairment effect observed for the 10-min group was not large. Also, the main-effect difference between test performances at 10 min and at 48 hr was large, indicating that there was a large amount of extraexperimental forgetting, $F(1, 92) = 14.82$. Thus, the moderate impairment effect and the strong extraexperimental-forgetting effect might have made it difficult for an existing recovery effect to appear as absolute spontaneous recovery in Experiment 1.

were identical to those of Experiment 1. Immediately after the first presentation of the interpolated faces, the experimenter introduced the second presentation. For the told-same participants the experimenter introduced the second presentation by saying, "Now I am going to have you rate these people a third and last time. . . ." For the told-different participants, the experimenter said, "Now I am going to have you rate, for a second and last time, the people you just saw. . . ." All participants were told to rate how likable each individual appeared by marking a continuum line anchored by *not very likable* and *very likable*. Participants then viewed the faces at a pace of 8 s per face.

After their second viewing of interpolated faces, all participants spent 5 min completing a questionnaire that was unrelated to the present study. Participants in the 10-min group then read cartoons for 5 min before starting the memory test. Participants in the 45-min group read cartoons for 5 min, watched a 35-min videotape of animated cartoons, and then started the memory test. Participants in the 48-hr group were dismissed until the second session, at which they read cartoons for 5 min before starting the memory test. After finishing the memory test, all participants were given a questionnaire. The questionnaire asked participants to recall what their initial thoughts were upon finishing the honesty rating task. Participants indicated whether they had thought the faces they rated on honesty were: (a) faces that they had already rated on attractiveness, (b) new faces that they had not rated on attractiveness, or (c) a mix of faces that they had rated on attractiveness, as well as new faces that they had not rated on attractiveness. Participants responding with the third option also indicated what percentage of the faces they thought were new ones.

Results and Discussion

The mean control scores, experimental scores, and memory impairment scores are displayed by groups in Table 2. The control and experimental scores were submitted to a 2 (told-same/told-different) × 3 (delay) × 2 (control/experimental) ANOVA. Overall, the participants correctly identified 73.0% of the control-condition targets (M = 16.1 of 22, SEM = 0.19) but only 66.8% of the experimental condition targets (M = 14.7 of 22, SEM = 0.22). The ANOVA indicated that this memory-impairment main effect was significant, $F(1, 138) = 39.13$, MSE = 3.48. As in Experiment 1, there was a significant Delay × Control/Experimental interaction, $F(2, 138) = 10.11$, MSE = 3.48, indicating that the memory impairment effect differed in strength among the delay groups. In both the 10-min group and in the 45-min group, control scores (Ms = 16.5 and 16.4, SEMs = 0.33 and 0.33, respectively) were significantly higher than experimental scores (Ms = 15.1 and 13.9, SEMs = 0.32 and 0.46), $F(1, 138) = 14.67$, and $F(1, 138) = 44.56$, respectively (MSEs = 3.48). In the 48-hr group, however, the control scores (M = 15.2, SEM = 0.32) were nearly equivalent to the experimental scores (M = 15.1, SEM = 0.32), $F(1, 138) < 1$.[4] Thus, both the 10-min and 45-min groups exhibited robust memory impairment effects, but evidence of this impairment did not exist after a 48-hr delay.

As was the case for Experiment 1, there was a particular interest in whether a three-way interaction would suggest that only told-same participants, not told-different participants, experienced a release from memory impairment. The data did not show such a pattern, however. Although the three-way interaction was nearly significant, $F(2, 138) = 3.07$, MSE = 3.48, p = .05, this effect seems to be the result of an inexplicably low control-condition score from told-different participants in the 10-min group.[5] Within the 45-min and 48-hr groups, the Control/Experimental × Told-Same/Told-Different interactions were not significant, $F(1, 138) = 1.72$, and $F(1, 138) < 1$, respectively. Thus, there is no evidence to suggest that the told-different and told-same participants differed in the recoverability of the memory impairment they suffered. Both told-same and told-different participants experienced a form of memory impairment that persisted through 45 min but apparently subsided within 48 hr. The two groups of participants also did not differ on their overall test performances or their performances within delay groups; neither the told-same/told-different main effect nor the Told-Same/Told-Different × Delay interaction effect was significant, $F(1, 138) < 1$, and $F(2, 138) < 1$, respectively (MSEs = 8.10).

The results suggest that the interpolated faces caused both told-same and told-different participants to suffer a temporary form of memory impairment. This interpretation was also offered for the results of Experiment 1 but was accompanied by an alternative interpretation. According to the alternative interpretation, the relative spontaneous recovery effect observed in Experiment 1 was not due to a release from memory impairment. This interpretation, however, does not appear to

Table 2
Mean Control-Condition and Experimental-Condition Performance (in Percentages) by Groups in Experiment 2

Group	Control condition		Experimental condition		Impairment effect (%)
	M	*SEM*	*M*	*SEM*	
10-min delay					
Told-same	79.0	1.8	66.9	2.1	12.1
Told-different	71.4	2.1	70.3	1.9	1.1
45-min delay					
Told-same	75.4	2.1	61.6	3.6	13.8
Told-different	74.2	2.1	65.0	2.3	9.2
48-hr delay					
Told-same	68.6	2.2	68.6	2.2	0.0
Told-different	69.7	1.9	68.6	1.9	1.1

Note. The means are based on the percentages of the 22 targets correctly identified. The sample size for each mean is 24. The impairment-effect column indicates the mean differences between the control and experimental performances.

[4] The power to detect a memory impairment effect in the 48-hr group was about .81 (assuming a one-tailed alpha of .05 and using the effect size estimate from the 10-min group, d = .53).

[5] The told-different participants in the 10-min group had a control-condition performance of only 71.4%. I can think of no meaningful account for this low performance. One would expect that the control performance of told-same and told-different participants would be equivalent, but told-same participants in the 10-min group had a much higher control performance of 79.1%. The control performance of told-different participants in the 10-min group was also low relative to the control performances of the participants in the 10-min groups of Experiment 1. In addition to being the apparent source of the significant three-way interaction in Experiment 2, the low performance of participants in the 10-min, told-different group seems to be the source of a significant Control/Experimental × Told-Same/Told-Different interaction, $F(1, 138) = 5.76$, and a significant Control/Experimental × Told-Same/Told-Different interaction within the 10-min group, $F(1, 138) = 10.07$.

be applicable for the results of Experiment 2. Figure 2 provides a telling graphical representation with mean control and experimental scores plotted for each of the three delay groups. Unlike in Experiment 1, the main effect for the delay variable was not significant, $F(2, 138) = 1.65$. As expected, normal forgetting caused performance at recognizing control-condition faces to decline over the delay groups. This performance pattern, however, was not exhibited for experimental-condition faces. It appears that normal forgetting caused experimental-condition performance to drop from the 10-min delay group to the 45-min delay group, but performance showed no drop from the 45-min group to the 48-hr group. In fact, from the 45-min group to the 48-hr group, experimental-condition performance increased from 63.2% to 68.6%. This increase was small (5.4%) but significant, $F(1, 138) = 4.90$, $MSE = 6.66$. The pattern of results between the 45-min and 48-hr delay constitutes an absolute spontaneous recovery effect. (See Experiment 3 for a discussion of possible reasons for the small size of this absolute recovery effect.) As mentioned above, an absolute recovery effect indicates that there was a release from the observed memory impairment.

Release from impairment suggests that the impairment observed at the short retention intervals in the present experiment was retrieval-based rather than storage-based. Assuming that memories that are altered or replaced cannot be spontaneously unaltered or unreplaced, storage-based hypotheses cannot account for a release from memory impairment. Only retrieval-based hypotheses, which posit that memories for both the target and the interpolated face coexist in memory, can explain how a once-impaired memory could later become unimpaired. The results also suggest that told-same and told-different participants alike were plagued by retrieval-based impairment. Neither group exhibited any signs of storage-based impairment.

Manipulation Check

The finding that told-same participants experienced a temporary form of memory impairment seems contrary to the hypothesis that interpolated faces would cause storage-based impairment when those faces were thought to be the same as the target faces. It is possible, however, that the told-same participants, despite attempts to convince them otherwise, believed that the interpolated faces were different from the target faces. Responses to the question assessing the told-same/told-different manipulation were analyzed to investigate this possibility. Of the told-same participants in the 10-min group, the majority (58.3%) reported thinking that all of the interpolated faces had already been seen as target faces. On average, told-same participants in the 10-min group thought that only 10.8% ($SD = 16.2$) of the interpolated faces were new faces that were not seen as targets. These findings strongly suggest that the told-same participants accepted the misleading suggestions of the experimenter and that they adopted and maintained the general belief that the interpolated faces were already seen as targets. Surprisingly, told-different participants seemed somewhat less accepting of the information provided by the experimenter. Of told-different participants in the 10-min group, 20.8% reported thinking that all of the interpolated faces had already been seen as target faces, whereas 29.2% thought that all of the interpolated faces were new, previously unseen faces. On average, told-different participants in the 10-min group thought that 50.9% ($SD = 39.2$) of the interpolated faces were new faces that were not seen as targets. This mean percentage was significantly different from the mean of percentages reported by told-same participants in the 10-min group, $U(48) = 111.5, p < .001$.

Participants' self-reported thoughts about the interpolated faces indicate that told-same participants held the belief that the interpolated faces were the same as the target faces, whereas the beliefs of told-different participants showed more variability. These self-reports were used in internal analyses to investigate whether the beliefs of the participants, irrespective of whether they were in the told-same or told-different condition, might predict the strength or type of memory impairment they experienced. Spearman correlation coefficients were computed for the relations between the percent of interpolated faces that participants thought were new faces and the amount of memory impairment exhibited (computed by subtracting each participant's experimental score from control score). The overall coefficient ($r = -.04$), as well as the coefficients for the 10-min, 45-min, and 48-hr groups, was not significant at the .05 alpha level ($rs = -.25, -.05, .16$, respectively). A similar analysis was conducted by calculating correlation coefficients for the relation between the percent of interpolated faces that were thought to be new and the experimental scores, with control scores partialled out. Both types of analyses lead to the same conclusion: The magnitude of the memory impairment suffered by participants was not a function of the beliefs they held about the interpolated faces.

An Additional Manipulation Check

Because the told-same/told-different manipulation yielded null results on the recognition measure used in the present experiments, and because these null results have important theoretical and practical implications, it is important to give careful consideration to any concerns as to whether the

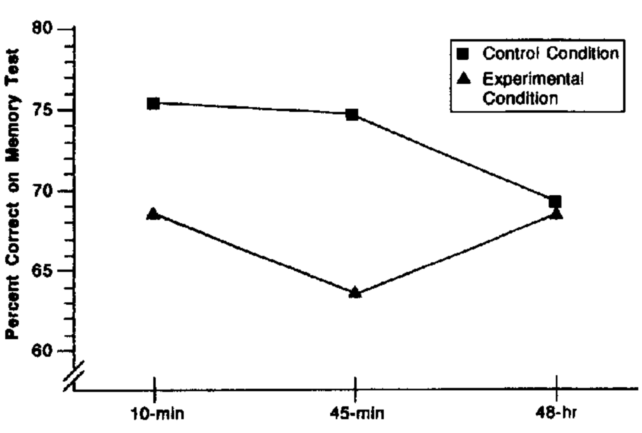

Figure 2. The percentage of control targets and experimental targets that were correctly identified as a function of delay group and collapsed across the told-same/told-different groups in Experiment 2.

manipulation instructions had the intended effect. Although the manipulation check used in Experiment 2 yielded robust differences between the told-same and told-different groups, it could be argued that the manipulation check was not a good indicator of participants' beliefs concerning the relationship between the interpolated and target faces. A reader of a previous version of this article suggested that participants' responses on the manipulation-check measure may have been biased by the told-same/told-different instructions, even though their actual beliefs regarding the relationships between interpolated and target faces were not influenced by those instructions. Participants might not have been affected by the instructions, but when they encountered the manipulation-check question, they recalled having been given information about the relationship between the interpolated and target faces.

It seems unlikely that this interpretation describes what occurred. If participants were able to recall the instructions given to them, they must have been aware of those instructions at the time they were given. It would be unusual if participants ignored those instructions while they viewed the interpolated faces but later utilized that information to answer the manipulation-check question. Also, most participants (61%) reported that they thought there was a mix of previously seen and unseen faces in the interpolated presentation. This fact suggests that participants did not simply give a response based on memory of the instructions they received (otherwise participants would have preferred "all" or "none" responses). It seems unlikely that the "mix" responses occurred because participants compromised between what their true beliefs were and what they remember about the instructions. This would require participants to have thoughts such as "I believe that all of the faces were new, but the experimenter said that I saw them all before, so I'll put 50% down for this question." It seems more likely that participants would report on their beliefs when specifically asked about them.

Data from another experiment provide further evidence that the told-same/told-different instructions had the intended effect and that participants' responses on the manipulation-check question were not simply recountings of what was remembered about the instructions. With two exceptions, the conditions and procedures experienced by participants in this experiment were identical to those experienced by participants in the 10-min groups of Experiment 2. One exception was that the participants in this experiment were given a different type of memory test, the results of which are not reported here. The exception of interest was that participants in this study received a different manipulation-check questionnaire. The questionnaire first informed participants that the experimenter flipped a coin to determine the type of instructions to give in the session (whether to indicate that the faces in the second presentation are the same or different from those seen in the first presentation). Participants were then asked to indicate whether they were given told-same or told-different instructions. For this question, 24 of 27 told-same participants, and 23 of 27 told-different participants gave correct answers. The questionnaire then informed participants that the experimenter blindly drew one of five slips of paper out of a box to determine which of five videotapes to show for the second presentation of faces. The questionnaire indicated that on one tape, 5% of the faces were different from those seen in the first presentation of faces, on another 25% were different, on another 50% were different, on another 75% were different, and on another 95% were different. Participants were then asked to indicate which tape they believed they saw. A large majority of told-same participants (21 of 27) picked tapes that were described as containing 25% or fewer different faces (13 participants picked the 5% tape and 8 picked the 25% tape). For the told-different participants, a large majority (19 of 27) picked tapes that were described as containing 50% or more different faces (9 participants picked the 50% tape, 5 picked the 75% tape, and 5 picked the 95% tape). These manipulation-check results are a bit more convincing regarding the effectiveness of the told-same/told-different manipulation than the manipulation-check results of Experiment 2. The questionnaire containing the manipulation-check question made participants aware that there was no reason for them to rely on the experimenter's instructions in deciding how to answer the manipulation-check question. This rules out the possibility that participants' beliefs were unaffected by the told-same/told-different instructions and that participants were simply reporting on their memory for the instructions.[6]

Experiment 3

Although the results of Experiment 2 included a statistically significant absolute spontaneous recovery effect, the effect was surprisingly small. It is difficult to build full support for conclusions that are based on such an effect. Experiment 3 was conducted for purposes of collecting additional evidence that the impairment detected in the present experiments was temporary and that even participants in a told-same group experienced a release from impairment.

As mentioned above, there are two methods for demonstrating a release from impairment. One is to detect an absolute spontaneous recovery effect, and a second involves investigating the location of floor-level performance. To investigate whether the absolute recovery effect of Experiment 2 would replicate, Experiment 3 included conditions identical to the 45-min and 48-hr groups of Experiment 2. To investigate the plausibility of floor-level explanations for the recovery effects observed in the present experiments (1, 2, & 3), a 1-week group was also included. It was expected that performances in the 1-week group would be significantly poorer than performances in the 45-min group and, more importantly, in the 48-hr group. This result would indicate that the disappearance of the memory impairment effect at 48 hr could not be explained by assuming that control and experimental perfor-

[6] One note about the manipulation-check measure seems warranted. Although the manipulation-check results presented are effective in demonstrating a difference in the beliefs of told-same and told-different participants, they might underestimate the effectiveness of the manipulation instructions used in the experiments. It seems quite possible that most participants fully accepted the instructions and, while viewing the interpolated presentations, held the "appropriate" beliefs regarding the interpolated faces. However, only when these participants were confronted with the manipulation-check question did they begin to suspect that the instructions and their beliefs were not completely correct.

mances reached a floor level at 48 hr. This result would provide unambiguous evidence that participants experienced a release from the impairment detected at 45 min.

Method

Participants and Design

The participants were another 192 students from the same pool used in Experiments 1 and 2. The design was a 3 (45-min, 48-hr, 1-week) × 2 (Control/Experimental) factorial. Unlike in Experiments 1 and 2, there was no told-same/told-different variable.

Materials and Procedure

The materials and procedures used in the experiment were identical to those used in Experiment 2, except for the following three changes. All participants were given the told-same instructions used in Experiment 2. Participants in the 1-week group experienced the same materials and procedures (not delay) as did the participants in the 48-hr group. The measure used to assess the effectiveness of the told-same instructions was identical to the manipulation-check measure described above (under the heading *An Additional Manipulation Check* in the *Results and Discussion* section for Experiment 2).

Results and Discussion

The mean control scores, experimental scores, and memory impairment scores are displayed by groups in Table 3. The control and experimental scores were submitted to a 3 (Delay) × 2 (Control/Experimental) ANOVA. Overall, the participants correctly identified 71.0% of the control-condition targets (M = 15.6 of 22, SEM = 0.17) and 67.6% of the experimental condition targets (M = 14.9 of 22, SEM = 0.17). This memory-impairment main effect was significant, $F(1, 189) = 12.41$, $MSE = 4.29$. As in Experiments 1 and 2, there was a significant Delay × Control/Experimental interaction, $F(2, 189) = 6.34$, $MSE = 4.29$, indicating that the memory-impairment effect differed in strength among the delay groups. In the 45-min group, the memory-impairment effect (i.e., the difference between control and experimental scores) was significant, $F(1, 189) = 24.08$, $MSE = 4.29$. In the 48-hr group and 1-week group, however, there was no evidence of memory impairment effects, $F(1, 189) < 1$ for both comparisons.[7]

How can the disappearance of the memory impairment effects after 48 hr be explained? An explanation involving floor-level performance cannot account for why participants in the 48-hr group exhibited no impairment. The overall performance of participants in the 1-week group (65.3%) was substantially lower than the performance of participants in the 48-hr group (69.7%), $F(1, 189) = 9.03$, $MSE = 6.44$.[8] Therefore, the functional floor level must lie significantly below the level of performance exhibited by participants in the 48-hr group. Figure 3 helps illustrate why one cannot explain the disappearance of memory impairment at 48 hr by assuming that experimental and control performance had hit a functional floor level. Had experimental and control performance reached a floor level, there would be no decrease in performance between the 48-hr and 1-week groups. These results indicate that the impairment suffered by participants after 45 min was temporary and that a release from impairment occurred before the end of the 48-hr delay.

Unlike in Experiment 2, a significant absolute spontaneous recovery effect was not detected. The difference between experimental-condition performances in the 45-min and 48-hr groups was directional, but trivial $F(1, 189) < 1$. The power to detect an absolute recovery effect in the present experiment (given the size of the effect in Experiment 2, a one-tailed alpha level of .05, and a sample size of 64 per group) was reasonably high (about .80), and therefore, the failure to replicate the effect found in Experiment 2 was somewhat surprising. However, the fact that a relative recovery effect was detected, rather than an absolute recovery effect, should not be taken as evidence against the idea that recovery occurred. The results concerning the location of the floor-level performance clearly indicate that the relative recovery effect must be attributed to a release from impairment. The relative recovery effect indicates that the amount of recovery occurring between 45 min and 48 hr was roughly equal to the amount of extraexperimental forgetting that occurred in that time period.

In hindsight, one can offer an account for the small absolute recovery effect detected in Experiment 2 and the difficulty in replicating it. The stimuli used in the present experiments, namely faces, are not-at-all unique to an experimental setting. The people who participated in these studies see hundreds of faces each day (e.g., in class, on TV, etc.). For people participating in a 45-min delay condition, no real-world faces, except the faces of the experimenter and fellow participants, intervene between the target presentation and the memory test. For people participating in a 48-hr delay condition, however, hundreds of faces intervene between the target presentation and the memory test; these participants leave the laboratory for 48 hr. Their exposure to extraexperimental faces may be a significant source of extraexperimental forgetting. Although a substantial amount of recovery occurred between 45 min and 48 hr, so did a substantial amount of extraexperimental forgetting. Consequently, relative recovery effects were

Table 3
Mean Control-Condition and Experimental-Condition Performance (in Percentages) by Groups in Experiment 3

Group	Control condition *M*	Control condition *SEM*	Experimental condition *M*	Experimental condition *SEM*	Impairment effect (%)
45-min delay	76.8	1.4	68.6	1.5	8.2
48-hr delay	70.5	1.3	68.9	1.2	1.6
1-week delay	65.6	1.2	65.2	1.4	0.4

Note. The means are based on the percentages of the 22 targets correctly identified. The sample size for each mean is 64. The impairment-effect column indicates the mean differences between the control and experimental performances.

[7] The power to detect a memory impairment effect in the 48-hr or 1-week groups was about .96 (assuming a one-tailed alpha of .05 and using the effect size estimate from the 45-min group, d = .61).

[8] It was also the case that the experimental performance after the 1-week delay was significantly poorer than experimental performance after the shorter delays (45 min and 48 hr), $F(1, 189) = 4.52$. The main effect for delay was also significant, $F(2, 189) = 13.01, p < .001$.

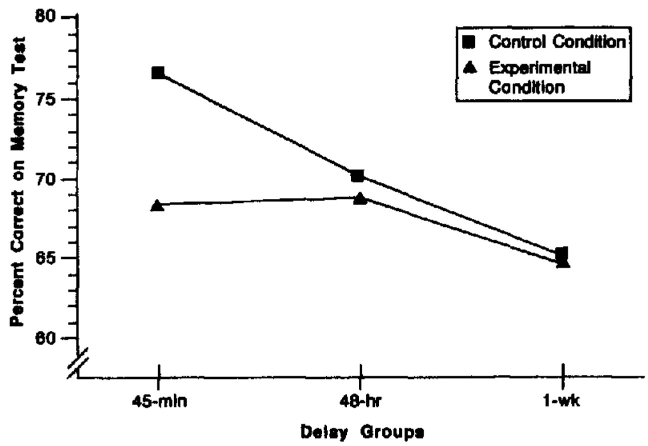

Figure 3. The percentage of control targets and experimental targets that were correctly identified as a function of delay group in Experiment 3.

detected rather than absolute recovery effects. This idea suggests that absolute spontaneous recovery effects might best be investigated by using stimuli that are not overly common in participants' everyday experiences, or by restricting participants' extraexperimental exposure from items similar to the critical stimuli.

All participants in the present experiment were given the told-same instructions used in Experiment 2. One of the motivations for these experiments was to determine whether storage-based impairment would occur when a person believed that the interpolated and target faces were the same. For reasons discussed above, it is evident that the impairment detected for participants in the present experiments was not storage based. Although data presented in the *Results and Discussion* section of Experiment 2 indicated that told-same and told-different instructions used in the present experiments were effective, additional data were collected in Experiment 3 to ensure that the told-same instructions had the intended effect on participants' beliefs. Upon completion of the memory test, participants were given a measure that was identical to the manipulation-check measure described in the *Results and Discussion* section of Experiment 2. Recall that this measure asked participants to guess which of five interpolated video tapes (described as containing 5%, 25%, 50%, 75%, or 95% new faces) they saw. The results support the efficacy of the told-same instructions. Of participants in the 45-min group, a large majority of participants (44 of 64) picked tapes described as containing 25% or fewer different faces (23 participants picked the 5% tape and 21 picked the 25% tape). Only 5 participants picked the 75% tape, and only 1 picked the 95% tape. The results of internal analyses indicated that even participants who were completely misled regarding the interpolated faces experienced a release from memory impairment between the 45-min and 48-hr delay. In these analyses, only participants who picked the 5% tape were included ($N = 71$). The overall pattern of means for these people was essentially identical to the overall pattern of means shown in Figure 3.

Experiment 4

The results of Experiments 1, 2, and 3 demonstrate that memories for faces are susceptible to impairment. A natural conclusion regarding the impairment would be that participants' memories for a target face were impaired because they viewed the interpolated face to which it was related. This conclusion, however, might be premature. When participants viewed interpolated faces, they also viewed the labels (e.g., *Bride*) accompanying those faces. Thus, participants saw the labels for experimental-condition targets more often than for the control-condition targets. A reader of a previous version of this article suggested that this alone might account for differences in experimental-condition and control-condition test performance. Experiment 4 was conducted to test this possibility.

A possible explanation for how exposure to labels themselves can cause impairment involves the phenomenon of retrieval-induced forgetting. Findings from experiments investigating retrieval-induced forgetting suggest that active retrieval of information related to a cue can impair retrieval of other information related to that cue (see Anderson, Bjork, & Bjork, 1994; Blaxton & Neely, 1983; Brown, 1981). In the present experiments, a form of this phenomenon might be affecting participants' abilities to retrieve stored memories for experimental-condition target faces. Exposure to labels in the interpolated presentation may cause participants to retrieve images of various people who might fit that label, which in turn produces retrieval-induced forgetting for the target that had accompanied that label. Thus, the cause of the impairment observed in the present studies might not have been participants' exposure to interpolated faces, but rather their exposure to retrieval processes that were triggered by interpolated labels.

To test the interpretations involving exposure to interpolated labels, Experiment 4 utilized a fully within-subjects design that included four conditions—control, experimental, label only, and repetition. The control and experimental conditions were analogous to those used in the previous experiments—target faces in the control condition were not followed by related faces in the interpolated presentation, whereas target faces in the experimental condition were followed by related interpolated faces that shared the same label. In the label-only condition, labels that had accompanied target faces appeared in isolation in the interpolated presentation. For example, a participant may have seen the labeled face of a bride in the target presentation, but saw only the label *Bride* in the interpolated presentation. In the repetition condition, the same labeled faces that were seen in the target presentation were seen again in the interpolated presentation.

As in Experiments 2 and 3, participants in Experiment 4 made attractiveness ratings for the target presentation and honesty and likableness ratings for the interpolated presentation. In instructions about the interpolated presentation, participants were forewarned that they might see displays of labels in isolation, and they were instructed to make an honesty rating for each of those displays based on their perceptions of how honest the typical person who fits the presented label is. If participants have a tendency to retrieve images of persons in response to interpolated labels, that tendency should be enhanced by these instructions regarding the label-only displays. Therefore, if the impairment effects observed in the present experiments are the result of retrieval-induced forgetting prompted by exposure to interpolated labels, there should be a significant amount of memory impairment observed for the label-only condition.

Method

Participants and Design

The participants were another 32 students from the same pool used in Experiments 1, 2, and 3. A four-condition within-subjects design was used.

Materials and Procedure

The materials and procedures used in the experiment were very similar to those used in Experiments 2 and 3, but with some exceptions. As in Experiments 2 and 3, all participants viewed and made attractiveness ratings for 8 introductory faces and 44 target faces contained in a target videotape presentation. After the target presentation, participants read cartoons for 2 min. The experimenter then introduced the interpolated presentation by saying the following:

> Now you'll be seeing more portraits, but this time you should rate how honest you think the person in the portrait appears. There will also be instances in which just labels appear, without a face. When that happens, you should still make a rating, but you should base your rating on how honest you think the typical person who fits that label is.

Each interpolated presentation contained an introductory segment comprised of three labeled faces that had already been seen as introductory faces in the target presentation and one instance in which a label that had accompanied an introductory face in the target presentation was displayed in isolation. The remainder of the interpolated presentation contained 11 labeled faces that were related to target faces seen earlier (experimental condition), 11 labeled faces that were already seen as target faces (repetition condition), and 11 instances in which a label that had accompanied a target face was displayed in isolation (label-only condition). For the 11 target faces from the control condition, no faces or labels were displayed in the interpolated presentation.

After making honesty ratings for the faces and labels in the interpolated presentation, participants viewed the presentation again and made ratings of likableness. A 5-min delay period followed in which participants read cartoons. Participants' memories for target faces were then tested with a modified recognition test, identical to those used in the previous experiments.

The randomization, counterbalancing, and ordering schemes used in Experiment 4 were analogous to those used in the previous experiments, but with one exception. A given face was seen either as a target, interpolated, or novel face (not counterbalanced across the three, but rather, randomly assigned to one of the three). As in the previous experiments, however, the condition in which a given face served was fully counterbalanced.

Results and Discussion

Mean performances a correctly identifying target faces were 77.8% in the control condition (M = 8.6, SEM = 0.23), 70.5% in the experimental condition (M = 7.8, SEM = 0.30), 77.3% in the label-only condition (M = 8.5, SEM = 0.30), and 93.2% in the repetition condition (M = 10.3, SEM = 0.18). An ANOVA yielded a significant overall effect for these within-subjects conditions, $F(3, 93) = 28.28$, $MSE = 1.26$. Planned comparisons indicated that experimental-condition performance was significantly lower than control-condition performance, which replicates the memory impairment effects observed in the previous experiments, $F(1, 31) = 5.51$, $MSE = 1.91$. Performance in the label-only condition, however, was nearly equivalent to control-condition performance, which indicates that exposure to labels themselves produced no detectable impairment, $F(1, 31) < 1$. A direct comparison of the experimental and label-only conditions reveals a significant performance difference attributable to the impairment affecting the experimental-condition performance, $F(1, 31) = 6.80$, $MSE = 1.32$. Finally, and not surprisingly, performance in the repetition condition was significantly higher than performance in the control condition, $F(1, 31) = 49.67$, $MSE = 0.92$.

The results indicate that the memory impairment effects observed in the present experiments cannot be attributed to the fact that labels for the experimental targets were viewed more often than labels for the control targets. The interpolation of labels in the label-only condition produced no detectable impairment, whereas the interpolation of labeled faces in the experimental condition produced robust levels of impairment. These findings rule out the retrieval-induced forgetting interpretation for the memory impairment effects and indicate that the effects observed in the three experiments were caused by the interpolation of faces related to the targets.

General Discussion

There were three main findings of the present experiments. First, interpolated faces impaired participants' abilities at recognizing previously seen target faces. In all four experiments, robust memory impairment effects were detected for participants tested after short retention intervals.[9] These effects provide clear evidence that memories for faces are susceptible to impairment. The second finding was that the impairment caused by the interpolated faces was temporary. Spontaneous recovery effects were observed in Experiments 1, 2, and 3, and Experiments 2 and 3 provided specific evidence for a release from the observed memory impairment. The temporary nature of the impairment indicates that the interpolated faces did not alter or otherwise affect the storage of the target memories, but rather affected the retrievability of those memories. The third main finding was that the spontaneous recovery from impairment was observed for participants who believed that the target and interpolated faces were different, as well as for participants who believed the target and interpolated faces were the same. These three findings are discussed in more detail below.

Impairment of Memories for Faces

Many of the memory-for-faces studies that have examined the effects of interpolated faces were initially designed for

[9] I should note, however, the inexplicable absence of an impairment effect for told-different participants in the 10-min group of Experiment 2 (see footnote 5).

applied purposes. Researchers conducting the studies were concerned with questions such as how an eyewitness's lineup selection might be affected by misleading information, or how an interpolated mug-shot search can affect the likelihood with which an eyewitness will make a positive identification or false identification. Numerous such studies have produced findings that suggest that interpolated information can have important effects on relevant behaviors of an eyewitness. For example, Loftus and Greene (1980) found that misleading postevent information can affect participants' choices on a memory test for a previously seen target person.

One question that has not been addressed in these studies, however, is whether recognition of a face can actually be impaired by exposure to interpolated information or an interpolated face. The studies showing that interpolated information or an interpolated face can affect subsequent test performances were not designed to specifically test for the occurrence of memory impairment. Therefore, the results from these studies are attributable to factors such as source misattributions, criterion shifting, demand factors, or other factors distinct from memory impairment. For example, the memory test used in Loftus and Greene's (1980) study included the postevent information as a test alternative. Participants in that study might have tended to select the interpolated information on the test because of perceived social demand or because they failed to encode the target information but did remember the interpolated information. As described earlier, the results of two studies demonstrating that mug-shot searches affected participants' subsequent identifications of targets can be attributed to criterion shifting (Laughery et al., 1971; Laughery et al., 1974). Participants in those studies may have become more reluctant to make any identification after searching through mug shots.

In the present experiments, the modified testing procedure was used to control for alternative interpretations and to measure memory impairment. The modified test excluded interpolated faces as response alternatives, thereby removing the possibility that response biases and perceived social demand could account for the performance effects caused by interpolated information. Instructions given to participants to select the test face seen at any time during the experiment removed problems with source misattributions. Also, the forced-choice nature of the modified test eliminated concern over criterion shifting. With alternative interpretations controlled for, the modified test used in the present experiments detected robust levels of memory impairment. Thus, these experiments provide the first demonstrations that memory for a face can be impaired by exposure to a related interpolated face.

Some previous memory-for-faces studies that had controlled for certain alternative interpretations did not find similar evidence that interpolated faces can impair recognition performance (see Davies et al., 1979; Forbes as cited in Davies et al.) Davies et al. found that participants who had rated the pleasantness of 100 interpolated faces were not significantly less accurate at identifying three previously seen target faces than were participants who did not see the 100 interpolated faces. Similarly, in an eyewitness identification study by Brigham et al. (1988), participants who rated the attractiveness of 18 interpolated faces were not significantly less accurate at later identifying a target individual than were participants who were not exposed to the interpolated faces.

It is not fully clear why Davies et al. (1979) and Brigham et al. (1988) did not observe impairment like that in the present experiments. The most straightforward of several possible reasons is that these studies lacked sufficient power to detect existing impairment. The sample sizes in both studies were small, and between-subjects designs were used. The sample sizes summed across the relevant conditions cited above were 30 in the Davies et al. study and 51 in the Brigham et al. study. In both studies, participants who had not been exposed to interpolated faces were more accurate at identifying targets than were participants not exposed to interpolated faces, but the performance differences were not significant. The average effect-size estimate (Cohen's d) for the memory impairment effects in the 10-min groups of the present experiments was about .50. Using this value to estimate power in the Davies et al. (1979) study, which had 15 participants per condition, one finds that the power to detect an effect of that size at the .05 alpha level (one-tailed) was only .38. For the Brigham et al. (1988) study, which had 25 or 26 participants per condition, power was only .54.

Another possible reason for the difference in results regarding memory impairment in the present experiments and those of Davies et al. (1979) and Brigham et al. (1988) concerns the fact that the interpolated faces in the previous studies shared no unique similarity to, or association with, the target faces. For the present experiments, one of the admittedly subjective criteria used to construct the face sets was that the three faces of each set be somewhat similar in appearance. The fact that some of the told-different participants mistakenly concluded that many of the target and interpolated faces were the same, and the fact that told-same participants were easily misled to believe they were the same, attests to the idea that the matched faces were similar. Similarity between the interpolated and targets faces might be a necessary condition for impairment to occur (McGeoch & McDonald, 1931), and this condition might not have been fully met in the Davies et al. and Brigham et al. studies. This reasoning, however, is highly speculative. The level of similarity between target and interpolated faces was not manipulated in those studies, and it was not manipulated independently of label type in the present experiments.

Less speculative is the idea that the association between a target and interpolated face was an important component of the impairment detected in the present experiments. Given the design and results of the experiments, it is clear that an interpolated face tended to cause greater impairment to a participant's memory for the target to which it was related (whether explicitly by label or implicitly by similarity) than to targets to which it was not related. Had an interpolated face caused as much impairment to unrelated targets as to its related target, the experimental-condition performance and control-condition performance would have been equivalent. In the studies of Davies et al. (1979) and Brigham et al. (1988), there were no specific associations made between the interpolated and target faces. Thus, the conditions of those studies did not favor the item-specific impairment observed in the present experiments.

The Magnitude of the Impairment

How severe is impairment caused by interpolated faces? The differences between the percentage of control targets and the percentage of experimental targets correctly identified was 4.0% for the 10-min group of Experiment 1, 6.6% for the 10-min group of Experiment 2, 11.5% for the 45-min group of Experiment 2, 8.2% for the 45-min group of Experiment 3, and 7.3% for the participants in Experiment 4 (which had a 5-min delay). Although these numbers may appear to reflect only moderate memory impairment effects, there are at least three reasons why they might underestimate the impairing effects of interpolated faces. The first reason is that the impairment measured in the present experiments was item specific; the reported impairment effects reflect the sum of the impairment that each interpolated face caused to the memory for the specific target to which it was matched. The impairment caused by other target faces or by unrelated interpolated faces is not reflected in the reported impairment effects. Thus, the impairment effects of Experiment 1, in which the interpolated faces were presented only once, demonstrate that memory for a given face can be impaired by only one exposure to a related face.

A second reason why the results might underestimate the impairing effects of interpolated faces concerns the fact that the modified test lacks sensitivity to certain forms of memory impairment (see Belli, 1989; Belli et al., 1992; Chandler, 1989; Loftus, Schooler, & Wagenaar, 1985; Tversky & Tuchin, 1989). For example, if a memory is only partially altered by an interpolated face, it might nevertheless, lead to a correct response on the modified test if cues to recognition of the target face remained unaffected. Also, impairment caused by interpolated faces might become more evident if the interpolated faces were offered as test alternatives (see Belli, 1989), but including the interpolated faces would also allow for alternative interpretations of the results.

The third reason why the results might underestimate the impairing effects of interpolated faces involves participants' considerable ability to guess correctly on the modified test. Because items on the modified test offer only two response alternatives, a participant who does not remember a target (for any reason, including impairment) has a 50% chance of guessing correctly at test. The high success rate for guessing makes the observed difference between the control and experimental scores an underestimate of the percentage of memories for experimental targets that were impaired. Consider the hypothetical case in which participants remember 100% of the control targets, but impairment from interpolated faces causes them to remember only 66.6% of the experimental targets. Given that participants would be 50% accurate when guessing about the nonremembered target faces, the resulting experimental-condition performance would be 83.3%. Hence, although 33.3% of the memories for experimental targets were impaired, the observed difference between control and experimental performance would be only 16.7%.

To estimate the proportion of memories for experimental targets that were impaired in the present experiments, the equation $c = r + .5(1 - r)$, can first be used to solve for the proportion of control targets and proportion of experimental targets that were remembered at the time of the test (Belli et al., 1992). The proportion of test items answered correctly (c) is a function of the proportion of targets remembered (r) and one half of the proportion of targets not remembered ($.5[1 - r]$). Solving for the proportion of targets remembered, results in $r = 2c - 1$.

For the 10-min group of Experiment 1, the proportion of control items that were answered correctly was .791, and the proportion of experimental items answered correctly was .751. By using the above equation it is found that, at test, participants remembered 58.2% of the control targets and 50.2% of the experimental targets. The difference between these percentages (8.0%) over the percentage of control targets that were remembered estimates the proportion of memories for experimental targets that were impaired (.137). That is, for about 14% of the experimental-condition faces that would have been remembered, impairment caused them to be not remembered. Calculations on the results from the Experiment 2 indicate that about 26% of the memories for experimental targets were impaired for the 10-min group and about 47% were impaired for the 45-min group. These numbers are more informative regarding the magnitude of memory impairment than are numbers reflecting the differences between control and experimental scores, and these numbers indicate that interpolated faces had substantial effects on recognition of target faces.

Impairment and Spontaneous Recovery

The memory impairment demonstrated in the present experiments was temporary. Spontaneous recovery effects were observed in Experiments 1, 2, and 3, and Experiments 2 and 3 provided specific evidence for a release from the observed memory impairment. These experiments, in which faces were used as stimuli, replicate those of Chandler (1991). In several experiments, Chandler found that interpolated nature scenes impaired participants' recognition of previously seen target scenes. As in the present experiments, the impairment was found to be temporary; spontaneous recovery was observed before the end of a 48-hr delay.[10]

The present demonstrations of spontaneous recovery (and those of Chandler, 1991) are different in several respects from previous recovery effects demonstrated in paired-associate experiments and the more recent recovery effects demonstrated by Wheeler (1995). First, the present experiments utilized a mixed-list design instead of the list-general designs used in most paired-associate experiments and in the experiments by Wheeler. Thus, the impairment and recovery in the present experiments must have been due to item-specific causes, and hypotheses assuming list-general designs (e.g., response-set suppression) are not candidates for explaining the observed impairment. Second, the present experiments

[10] Chandler (1991) does not report analyses for a direct comparison between experimental performances at different retention intervals. However, the experimental performances of participants in experiments with long retention intervals (e.g., 48 hr) appear to be greater than the experimental performances of participants in experiments with short retention intervals (e.g., 15 min). Hence, one could describe Chandler's experiments as having demonstrated absolute spontaneous recovery.

used stimuli (faces) and procedures (e.g., instructions promoting incidental learning) that were radically different from those used in paired-associate experiments or in Wheeler's experiments. In the typical paired-associate study, participants are explicitly instructed to learn word pairs, and in Wheeler's experiments, participants were explicitly instructed to remember sets of pictured objects, sets of words, or letter–word pairs. Third, and most importantly, the present experiments used a testing procedure that isolated the question of whether the interference effect (detected at the short retention interval) was due specifically to memory impairment.

Typical paired-associate experiments and the experiments by Wheeler (1995) involved memory tests that were sensitive not only to memory impairment, but also to other factors such as source misattributions. Thus, in Wheeler's experiments, for instance, it is not clear that the observed interference effects demonstrate memory impairment. Wheeler found that interpolated lists of items interfered with participants' recall of target-list items after a 1-min delay, but that this interference had lessened by the end of a 31-min delay. The test used to measure the interference was a recall test that was limited in duration (e.g., 1 min for Experiment 2). Correct performance required participants to make a discrimination regarding whether a recalled item was seen in a target list or an interpolated list. Because this discrimination effort would take time only for participants who were exposed to the interpolated lists (control participants did not see an interpolated list), a retroactive interference effect may have been observed even though memories for items on the target list were not impaired. Thus, although Wheeler's experiments demonstrated spontaneous recovery, they may have demonstrated recovery from interference related to source discrimination problems rather than for specific memory impairment. In the present experiments, however, the interference effects reflect impairment, and hence, the experiments are somewhat unique in providing a clear demonstration of recovery from impairment (in contrast to recovery from source confusion or other causes of interference effects).

Which Impairment Hypotheses Can Account for the Recovery?

The fact that there was recovery from impairment suggests that an entire class of hypotheses, the storage-based hypotheses, should be ruled out as explanations for the impairment that was observed after the 10-min delays. Had the impairment been caused by storage-based processes such as substitution or alteration (Loftus 1975, 1981a), that impairment would have persisted through the 48-hr delays.

It also seems that CHARM (Metcalfe, 1990), which relies on a storage-based assumption, would not do well in accounting for the present results. In CHARM simulations, it is assumed that when postevent information is encountered, a composite memory trace is formed that also includes the event information. In other words, event and postevent information are blended together (in a sense) and stored as one memory trace. CHARM has been used to successfully model the misinformation effects found in studies using the standard recognition test, as well as the noneffects found in studies using the modified recognition test (see Metcalfe, 1990, for details). It is unclear, however, how CHARM would model the recovery from impairment observed in the present experiments. If the experimental-condition targets and their related interpolated faces were stored as composite memories, and those composite memories caused the impairment effects detected after the 10-min and 45-min delays, then those same composite memories would have led to detectable impairment effects after the 48-hr delay.[11]

A small number of retrieval-based hypotheses of impairment can account for the recovery from impairment exhibited in the present experiments. The competition hypotheses suggest that when there are two coexisting memory traces for a given retrieval cue, the stronger, more recent memory can be accessed at the expense of the weaker one. This type of hypothesis can explain the present results by assuming that the recognition of a target face is affected by memory competition only when the memory of the related interpolated face is overwhelmingly strong (Chandler, 1991). After a delay of only 10 min, the memories for interpolated faces might be overwhelmingly strong as a result of their recency, but their advantage would have dwindled before the end of the 48-hr delay. One version of the competition hypotheses suggests that the retrieval of memories for old information is blocked when new information is recalled at test (Niemi, 1979). After the 10-min delays in the present experiments, memories for target faces may have been impaired through blocking because the interpolated faces were easily recalled. After the 48-hr delays, however, the interpolated faces were less likely to be recalled; thus, memories for target faces were unimpaired.

Although competition hypotheses such as the blocking hypothesis can account for the pattern of the present results, recent research has placed some doubt as to whether the mechanisms posited by such hypotheses were the cause of the observed impairment (Belli, 1993; Chandler & Gargano, 1995). For example, Chandler and Gargano (1995) found that when an interpolated nature scene (Lake B) was presented just before a modified test item (offering Target Lake A and Lake C as alternatives), no memory impairment was produced. Given that these conditions were optimal for observing impairment due to blocking, the absence of impairment makes the blocking hypothesis a poor explanation of impairment in related experiments.

Discrimination hypotheses suggest that impairment can result when two memory traces that share similar features are accessed simultaneously. There may be difficulty in discriminating between the memories, or the features may converge to

[11] The absence of an interference effect on the modified test does not indicate that no storage-based impairment occurred. Metcalfe (1990) has provided demonstrations of how an old memory that has been altered to accommodate new information (through blending) might nevertheless yield correct responding on the modified test. However, when an impairment effect that was once detected with the modified test is later absent, this suggests that the impairment originally detected was not storage based. There does not appear to be an account for how storage-based impairment, such as blending, would be detectable at one retention interval but completely undetectable at another.

form a composite memory (Reinitz, Lammers, & Cochran, 1992). A discrimination hypothesis involving convergence of retrieved features could account for the present results with the assumption that such a convergence occurred only when the memory for an interpolated face was adequately strong for retrieval. Memory impairment was not observed after the 48-hr delays because the memories for the interpolated faces had weakened beyond a certain critical point (Chandler, 1991). Other recent research has produced results that are favorably accounted for by discrimination hypotheses (see Chandler & Gargano, 1995).

There are other noteworthy impairment hypotheses that do not neatly fit into the storage/retrieval distinction used here. According to these hypotheses, which I refer to as inhibitory-based hypotheses, retrieval of target information is impaired when stored memory for target information is inhibited or suppressed (see Anderson & Bjork, 1994; Anderson & Spellman, 1995). Inhibitory-based hypotheses of impairment share characteristics with both the retrieval-based and storage-based hypotheses, as I describe them. For example, they are similar to retrieval-based hypotheses in that they do not involve threats to the integrity of the memory trace for old information. However, like for storage-based hypotheses, the processes causing impairment are thought to be initiated during an interpolated phase rather than at a final retrieval phase. Inhibitory-based hypotheses assume that the inhibition or suppression of a target memory can lift over time, and thus, they have no trouble in accounting for spontaneous recovery effects. However, can inhibitory-based hypotheses account for the existence of impairment in the present experiments?

Inhibitory-based hypotheses have been used to provide compelling explanations for various retrieval-induced forgetting effects. According to these accounts, retrieval of an item from memory (such as an interpolated response) is accompanied by the inhibition of related memory items (such as a related target response). Anderson et al. (1994) have recently suggested that retroactive interference effects found in paired-associate studies may have been retrieval-induced forgetting effects that are best explained by inhibitory-based hypotheses. In such studies the learning of interpolated (second-list) information often involved anticipation periods, in which a cue stimulus was presented and participants were to anticipate the correct second-list response before it was presented (see, e.g., Barnes and Underwood, 1959). Anderson et al. effectively argued that retrieval of second-list responses during these anticipation periods might have been accompanied by inhibition of memory for first-list responses, which partially or fully accounted for the retroactive interference observed in those studies. The arguments by Anderson et al. may be extended to any retroactive-interference study that requires participants to learn a new response while an old response competes; efforts to retrieve the new response would lead to the inhibition of the memory for the old response. In the present experiments, however, the interpolated phase did not require that participants actively retrieve interpolated faces. Rather, the faces were simply presented, and participants made no attempt to learn a new response and exclude an old one. There seems no reason for memories for target faces to have been inhibited while participants viewed the interpolated faces. It appears, therefore, that inhibitory-based hypotheses cannot account for the impairment observed in the present experiments.

Beliefs About New Information

As mentioned above, the spontaneous recovery effects demonstrated in the present research replicated effects found in experiments conducted by Chandler (1991) involving memories for nature scenes. There was a critical difference, however, between the present research and the work of Chandler. None of Chandler's participants were misled to believe that the interpolated scenes were the same as the target scenes. In the present experiments (1, 2, and 3), some of the participants were successfully misled to believe that the interpolated faces were the same as the target faces. The other participants were correctly informed that the interpolated and target faces were different. A surprising result occurred regarding this told-same/told-different manipulation—there was a null effect on the recognition measures.

As indicated in the *Results and Discussion* for Experiment 2, this null effect was not a result of a weak manipulation. Two independent manipulation check measures (as well as an additional measure from Experiment 3) indicated that the told-same participants accepted and maintained the belief that the interpolated faces were the same as the target faces. Although told-different participants were somewhat less accepting of the instructions given them, they on average reported believing that over half of the interpolated faces were different from the target faces. Internal analyses conducted for Experiment 2 indicated that participants who reported believing that the interpolated and target faces were different exhibited the same levels of impairment as participants who reported believing that the interpolated and target faces were the same.

There are at least two theoretical perspectives that would lead one to expect some type of difference in recognition performance between told-same and told-different participants. One perspective was developed by Hall et al. (1984), who suggested that memory impairment occurs only when discrepancies between postevent and event information are not detected (see also Tousignant et al., 1986). They reasoned that if a person detects discrepancies between the postevent and event information, the person would encode the postevent information in a distinct way such that it could be easily discriminated from original event information. In applying this hypothesis to the present experiments, one might have expected that told-different participants would not have experienced memory impairment (or would have experienced less than told-same participants) given that they were explicitly informed that there were discrepancies between the interpolated and target faces. A second perspective, developed in the introduction, suggested that when a person believes a new stimulus to be different from an old stimulus, the person's memory system creates a second and separate memory trace for the new stimulus, but the two coexisting memory traces would cause retrieval-based impairment. Alternatively, when a new stimulus is believed to be a second presentation of an already-stored memory, a memory system would modify the original memory to accommodate the new stimulus, resulting in storage-based impairment. In applying this hypothesis to the present experiments, it was expected that a temporary form of memory impairment would be observed for told-different participants, whereas the impairment suffered by told-same participants would be permanent and persist through the 48-hr delay.

Contrary to both of the hypotheses outlined above, the results of the experiments suggest that both told-same and told-different participants suffered equal levels of a retrieval-based form of impairment. These findings have important implications for understanding how new information affects memory for old information. Two of these implications are discussed below.

Discrepancy Detection and Impairment

One implication is that discrepancy detection and other factors that might cause people to believe that new information is different from old information do not preclude retrieval-based forms of memory impairment. This contrasts with Hall et al.'s (1984) idea that memory impairment does not occur when discrepancies are detected between postevent and event information. This implication also conflicts with certain evidence cited by Hall et al. (1984). For example, they discussed studies in which some participants were forewarned that some postevent information might be incorrect (see Greene, Flynn, & Loftus, 1982). In these studies, forewarned participants read the postevent information longer and were more accurate than nonforewarned participants at reporting critical event details. This difference in test accuracy was attributed to forewarned participants being more likely than nonforewarned participants to notice discrepancies between the event and postevent information and, therefore, being less susceptible to memory impairment caused by the postevent information. In these and other studies (e.g., Tousignant et al., 1986), however, there are reasons unrelated to memory impairment for why participants who noticed discrepancies would perform better on the memory test than participants who did not notice discrepancies. The memory tests used in such studies included postevent information as a response alternative. It could be the case that participants who noticed the discrepancies developed a strong memory for the postevent information and then selectively avoided postevent alternatives on the memory test. This would cause participants who noticed discrepancies to perform better than participants who did not notice discrepancies, even if the postevent information caused the same amount of memory impairment for both types of participants. In the present experiments, however, told-different participants (and participants who noticed discrepancies) had no advantage on the memory test, even if they developed a strong memory for interpolated faces; the interpolated faces did not appear on the test. With this control in place, the results indicate that even when discrepancies were announced to the participant, retrieval-based impairment occurred.

This finding suggests that there may be an underdeveloped approach to postevent information research. In most postevent information studies, the misinformation is inconspicuously embedded in a postevent presentation. The participants are led to believe that the postevent presentation is an accurate retelling of the event, and researchers hope that participants do not question the veracity of the postevent information. Although this approach is ecologically valid and does produce some consequential misinformation effects, the present results suggest that another approach, which also has ecological validity, might produce equally consequential memory-impairment effects. In such an approach, participants might be exposed to postevent information that is not framed as an accurate retelling. Perhaps its source could be a second eyewitness, a defense lawyer, a crime suspect, or a sketchy and preliminary news report. In such a study, participants probably would notice the discrepancies in the postevent information, but they may, nevertheless, suffer retrieval-based impairment because of it. Although a few postevent information studies have involved this type of postevent-information framing, the dependent variable of interest was memory reports (which can be influenced by social demand or by participants' trust in the postevent source) rather than participants' actual abilities to remember the event information (e.g., Jenkins & Davies, 1985; Loftus & Greene, 1980).

Conditions for Storage-Based Impairment

A second implication of the finding that told-same and told-different participants suffered retrieval-based impairment concerns the question of when storage-based forms of impairment might occur. Demonstrations of retrieval-based interference in retroactive-interference studies have not presented problems for proponents of storage-based impairment hypotheses. As discussed earlier, it seemed reasonable to argue that retrieval-based forms of impairment would occur when new and old information were believed to be different, whereas storage-based forms of impairment would occur when a person believed new and old information to be the same. Thus, prior to the present experiments, a proponent of storage-based hypotheses could have argued that the qualities of the memory impairment demonstrated in retroactive-interference experiments, such as Chandler's (1991), were not relevant to understanding the impairment produced in postevent information studies; the participants in Chandler's experiments and other retroactive-interference experiments were not misled, and therefore, those experiments were not conducive to storage-based forms of impairment. The results of the present experiments, however, undermine such arguments. The findings indicate that believing that new and old information is the same does not necessarily result in storage-based rather than retrieval-based impairment. They also suggest that although there may be important differences between the typical retroactive-interference study and the typical postevent information study, beliefs about the relationship between new and old information do not appear to be a sole critical distinction.

The results of the present experiments are consistent with the idea that long-term memory is permanent and not susceptible to storage-based impairment, and that retrieval failures are problems of accessibility, not availability (Shiffrin & Atkinson, 1969; Tulving, 1974). Nevertheless, it would be troublesome to suggest that the present results support such a conclusion. Rather, the present results may reflect the fact that some condition necessary for storage-based impairment to occur was missing from the experiments. Believing that new information is the same as old information may be a necessary, but not sufficient condition for new information to cause storage-based impairment. Assuming that is the case, what is the full set of conditions needed to produce storage-based impairment?

It is possible that storage-based impairment afflicts memories for only certain types of stimuli, such as objects or words but not faces. This seems unlikely, however. Faces are complex stimuli that are distinguishable by subtle and difficult-to-label differences in shapes, colors, and proportions. These qualities should make memories for faces excellent candidates for blending or other types of alteration.

Hall et al. (1984) have suggested that the activation of the memory for the old information might be necessary for storage-based impairment to occur. Only when a memory has been recently retrieved can it be altered. In this idea, there is some ambiguity as to what would constitute an adequate level of activation and why the level of activation would determine whether alteration occurred. Clearly, participants in the present experiments retrieved from memory the descriptive labels of target faces (e.g., *Bride*) upon seeing the same label on matched interpolated faces. Whether they recalled their visual memories of the target faces themselves is questionable. It is possible that memory alteration requires that participants recall visual memories for the targets while simultaneously looking at the interpolated faces.

There are several other factors that might potentially be included on a list of conditions necessary for storage-based impairment. Unfortunately, little work has been done to specify such a list. Perhaps researchers have shied away from distinguishing between conditions producing storage and retrieval impairment, given the limitations involved in demonstrating storage-based impairment. Showing that an erased memory was in fact erased, and that it will not sooner-or-later reappear, is impossible. Nevertheless, had the told-same participants in the present experiments exhibited impairment that persisted over time periods in which told-different participants experienced recovery, a relatively strong case for storage-based impairment could have been made. Thus, the present experiments provided a strong test of what was hypothesized to be a crucial determinant of whether storage-based impairment would occur. In doing so, the experiments also provided a methodology for further research that is greatly needed for clarifying issues regarding storage-based impairment.

Conclusion

This work began with two questions. First, are memories for faces susceptible to impairment? Previous face-memory studies had not fully isolated this question. The present experiments were the first memory-for-faces experiments to demonstrate impairment with a test that could rule out alternative interpretations involving demand factors, source misattributions, and other factors. The detected impairment was item-specific. This indicates that impairment for face memories is not restricted to instances in which a person is exposed to large numbers of new faces. Rather, it appears that memory for a face can be impaired by exposure to one related face. These findings have obvious practical significance regarding eyewitness memory. Also, they clearly refute the idea that memories for faces are somehow unique in their resistance to retroactive interference (see Davies et al., 1979; Davies & Christie, 1982), and they do not bode well for the claim that face recognition involves a special processing system (Whiteley & Warrington, 1977; Yin, 1969, 1970).

Second, how does new information impair memory for old information? The experiments tested the hypothesis that new information can cause either storage-based or retrieval-based impairment depending on whether the new information is believed to be the same or different from the old information. Despite successfully misleading participants in told-same conditions to believe that the interpolated faces were the same as the target faces, the impairment they exhibited was temporary, and thus, not storage-based. Robust levels of impairment and spontaneous recovery were observed regardless of what participants believed about the interpolated faces. The spontaneous recovery effects implicate retrieval-based hypotheses, such as a discrimination hypothesis, as being best suited to account for the detected impairment. The observation that participants experienced retrieval-based impairment regardless of their beliefs about the interpolated faces has at least two important implications. One implication is that, contrary to a common assumption in the postevent information paradigm, discrepancy detection between postevent and event information does not necessarily preclude memory impairment. A second implication is that believing that new and old information are the same is clearly not a sufficient condition for new information to cause storage-based impairment.

References

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). San Diego, CA: Academic Press.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58,* 95–105.

Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 139–145.

Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72–85.

Belli, R. F. (1993). Failure of interpolated tests in inducing memory impairment with final modified tests: Evidence unfavorable to the blocking hypothesis. *American Journal of Psychology, 106,* 407–427.

Belli, R. F., Windschitl, P. D., McCarthy, T. T., & Winfrey, S. E. (1992). Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 356–367.

Birnbaum, I. M. (1965). Long-term retention of first-list associations in the A–B, A–C paradigm. *Journal of Verbal Learning and Verbal Behavior, 4,* 515–520.

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from sementically related primes: Evidence of a category-specific inhibition. *Memory & Cognition, 11,* 500–510.

Brigham, J. C., & Cairns, D. L. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology, 18,* 1394–1410.

Brown, A. S. (1976). Spontaneous recovery in human learning. *Psychological Bulletin, 83,* 321–338.

Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 204–215.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116,* 38–49.

Chandler, C. C. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 256–265.

Chandler, C. C. (1991). How memory for an event is influenced by related events: Interference in modified recognition tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 115–125.

Chandler, C. C., & Gargano, G. T. (1995). *Retrieval processes that produce interference in modified forced-choice recognition tests.* Manuscript submitted for publication.

Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory & Cognition, 11,* 467–475.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987a). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72,* 629–637.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987b). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11,* 233–258.

Davies, G., & Christie, D. (1982). Face recall: An examination of some factors limiting composite production accuracy. *Journal of Applied Psychology, 67,* 103–109.

Davies, G., Shepherd, J., & Ellis, H. (1979). Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology, 64,* 232–237.

Deffenbacher, K. A. (1991). A maturing of research on the behavior of eyewitnesses. *Applied Cognitive Psychology, 5,* 337–402.

Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 299–305.

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21,* 207–219.

Hall, D. F., Loftus, E. F., & Tousignant, J. P. (1984). Postevent information and changes in recollection for a natural event. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 124–141). Cambridge, England: Cambridge University Press.

Jenkins, F., & Davies, G. (1985). Contamination of facial memory through exposure to misleading composite pictures. *Journal of Applied Psychology, 70,* 164–176.

Koppenaal, R. J. (1963). Time changes in strengths of A–B, A–C lists; spontaneous recovery? *Journal of Verbal Learning and Verbal Behavior, 2,* 310–319.

Laughery, K. R., Alexander, J. F., & Lane, A. B. (1971). Recognition of human faces: Effect of target exposure time, target position, pose position, and type of photograph. *Journal of Applied Psychology, 55,* 477–483.

Laughery, K. R., Fessler, P. K., Lenorovitz, D. R., & Yoblick, D. A. (1974). Time delay and similarity effects in facial recognition. *Journal of Applied Psychology, 59,* 490–496.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17,* 349–358.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7,* 560–572.

Loftus, E. F. (1981a). Mentalmorphosis: Alteration in memory produced by the mental bonding of new information to old. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 417–434). Hillsdale, NJ: Erlbaum.

Loftus, E. F. (1981b). Reconstructive memory processes in eyewitness testimony. In B. Sales (Ed.), *Perspectives in law and psychology: The trial process* (pp. 115–144). New York: Plenum Press.

Loftus, E. F. (1991). Made in memory: Distortions in recollection after misleading information. *Psychology of Learning and Motivation, 27,* 187–215.

Loftus, E. F., & Greene, E. (1980). Warning: Even memory for faces may be contagious. *Law and Human Behavior, 4,* 323–334.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35,* 409–420.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.

Loftus, E. F., Schooler, J. W., & Wagenaar, W. A. (1985). The fate of memory: Comment on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 114,* 375–380.

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

McGeoch, J. A. (1942). *The psychology of human learning.* New York: Longmans, Green.

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology, 43,* 579–588.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53,* 175–203.

Metcalfe, J. (1990). Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General, 119,* 145–160.

Niemi, P. (1979). A note on "blocking" as a retrieval interference phenomenon. *American Journal of Psychology, 92,* 547–550.

Postman, L. (1963). Does interference theory predict too much forgetting? *Journal of Verbal Learning and Verbal Behavior, 2,* 40–48.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7,* 672–694.

Reinitz, M. T., Lammers, L. J., & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20,* 1–11.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76,* 179–193.

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14,* 329–338.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62,* 74–82.

Tversky, B., & Tuchin, M. (1989). A reconciliation of the evidence on eyewitness testimony: Comments on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 118,* 86–91.

Underwood, B. J. (1948). "Spontaneous recovery" or verbal association. *Journal of Experimental Psychology, 38,* 429–439.

Wheeler, M. A. (1995). Improvement in recall over time without repeated testing: Spontaneous recovery revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 173–184.

Whiteley, A. M., & Warrington, E. K. (1977). Prosopagnosia: A clinical, psychological, and anatomical study of three patients. *Journal of Neurology, Neurosurgery, and Psychiatry, 40,* 395–403.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Yin, R. K. (1970). Face recognition by brain-injured patients: A dissociable ability? *Neuropsychology, 8,* 395–402.

Zaragoza, M. S. (1987). Memory, suggestibility, and eyewitness testimony in children and adults. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 53–78). New York: Springer-Verlag.

Zaragoza, M. S. (1991). Preschool children's susceptibility to memory impairment. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 27–39). Washington, DC: American Psychological Association.

Received January 11, 1995
Revision received February 16, 1996
Accepted February 16, 1996 ■