1994

The Influence Of Intervening Activities And Testing Conditions On The Accuracy And Confidence Of Source Memory

Joanne Louise Harbluk

ISBN 0-315-90576-X

ABSTRACT

The research in this thesis investigated factors that influence a particular type of source memory decision, the discrimination between memories that originate with others and memories that originate with ourselves. In their reality monitoring model, Johnson & Raye (1981) characterize the source decision process as an attribution based on an examination of memory rather than the retrieval of a label or tag that specifies a memory's source. Their research demonstrated that subjects rely on information from initial encoding to guide their decisions about the origins of their memories. The present research examined the influence of events occurring after encoding on source memory decisions. Postevent information (e.g., Loftus, 1991b) and the manner in which information is later processed (e.g., Jacoby, Kelley, & Dywan, 1989) are known to influence other types of memory judgments. More specifically, the purpose of the present research was to investigate the influence of events that occur after initial encoding, either during the period between study and test or at the time of testing, on the accuracy of source judgments and the confidence subjects express for those judgments. An additional focus of interest was the relationship between confidence in a memory and source decision accuracy.

The results of the present experiments indicated that source attribution errors were likely to occur when information was processed again in a manner that differed from that at original encoding. Source decision errors increased when the conflicting presentations were part of the actual memory test

(Experiments 1, 2, & 3) as well as when they occurred between study and test (Experiment 4). The confidence exhibited in source attributions was not always a reliable indicator of decision accuracy (Experiments 2 & 3) and under certain testing conditions subjects displayed inappropriate confidence for some of their erroneous source judgments. Finally, although ratings of memory strength (as indicated by confidence in recognition) generally predicted word recognition performance, strong memories were not necessarily associated with accurate source attributions (Experiments 3 & 4).

Overall, the results of the present studies are consistent with the reality monitoring model (Johnson & Raye, 1981) and the more encompassing source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993). Considered together, the results of the present studies extend the previous work by demonstrating that events occurring after a memory has been formed may influence decisions about how that memory was acquired.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge those people who contributed to the successful completion of this project. I would like to thank my advisor, Allan Paivio, and let him know that Omega appreciated his patient and consistent support throughout graduate school and especially with this research. I am also grateful to Albert Katz and Steve Lupker, my advisory committee, for their encouragement and many useful discussions; to Drs. Keith Humphrey, Ken Bowers, Bill Roberts, and Mark Kinnucan, my examiners, for their interesting comments and useful suggestions; and to Trudy Bons for her computing expertise.

Chris Roney, Randy Paterson, Sandra Houston, Natalie Allen, Carla Johnson, Mary Walsh, Deb Stuart, and Mary Anne Wilson provided friendship, humour, and some truly great dinners.

I would also like to express my gratitude to my family: my parents, June and Allan, for their unwavering belief that I could do it and the pride they have always shown in my accomplishments; Shirley and Jack, for the many ways they helped; and most of all Ian, for encouragement, steadfast support, and his gentle but skillful attempts at helping me keep things in perspective (many of which were successful).

Finally, this thesis is dedicated to my grandparents, Mina and Sam Herlovitch, who taught me how to play with words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

Introduction: The Problem of Memory Errors

Human memory is far from being completely reliable. We may forget where we have left our keys, fail to remember the name of a new acquaintance, or question whether we have turned off the stove once we have left the house. On those occasions when we are unable to come up with a desired bit of information, we become aware of the problem immediately. Memory, however, can fail us in other less obvious ways. For example, what we believe to be an accurate memory may bear little resemblance to the original event. It is only when we are confronted with evidence that contradicts our memory that we become aware of the memory distortion. If, however, that conflicting information never comes to light, the memory distortion may never be detected.

Although there is agreement that memory errors are a common experience, explanations differ as to why they occur. Freud studied many examples of memory errors and documented them in 'The Psychopathology of Everyday Life' (Freud, 1901/1972). Freud believed that memory errors were unconsciously motivated, a result of unresolved, underlying conflicts. Because he viewed these errors as systematic rather than random in nature, Freud believed that analysis of memory errors could provide 'windows to the mind.' Later, Bartlett (1932) also noted the nonrandom nature of memory errors. He observed that consistencies occur in memory distortions when people elaborate their recollections with what they believe to be true. For example, he found that details of a story based on events in another culture were altered to be more consistent with the rememberer's own culture. More recently, Reason and Mycielska (1982) reviewed experimental investigations of 'slips of the mind.' They examined many of the memory lapses which had intrigued Freud: slips of habit, recognition failures, memory blocks, losing track of past actions, repetitions and omissions, and lapses of attention and control. The explanations provided in these more recent investigations, however, account for memory errors in terms of the processes whereby the brain processes information - attention, perception, and cognition - rather than appealing to underlying psychological conflicts.

Memory errors are interesting and worth examining for a number of reasons. One primary motivation to do so is that memory errors are widely experienced and occur often enough to be considered part of normal human cognition. Although many memory errors result in annoyance and embarrassment, under some circumstances the consequences may be more serious. People generally consider their memories to be trustworthy. Our confidence in memory extends beyond our own recollections to the memories of others as demonstrated when eyewitness testimony is taken as persuasive evidence. While some memory errors may be recognized and corrected immediately, other 'memories' may be held with great conviction and acted upon as though they were genuine experiences. Our memories play an important role in defining our individual personal identity by providing the basis for our sense of self (Tulving, 1983). Johnson (1985) raises the interesting possibility that our view of who we are may be a composite of our true experiences, as well the products of our imagination which may include erroneous or misrepresented information. The potential threat to our self concept may in part account for our reluctance to challenge the veracity of our memories.

In summary, the study of memory errors can contribute to both practical and theoretical knowledge about memory. Among the practical considerations are issues of memory accuracy, the ability to monitor that accuracy, and individual differences in these tasks. An understanding of the relationship between memory accuracy and confidence is useful when evaluating memory reports as required, for example, in the case of legal testimony. In some working conditions the consequences of human error can be disastrous (Reason, 1990). Information about the conditions under which errors are most likely to occur could be used to design these environments to reduce the likelihood of errors.

The study of memory errors also contributes to theoretical knowledge about memory by providing information about the bases for our memory decisions. This work can illustrate, for example, the role that details of the original experience and reconstructive processes may play in memory. It can also provide clues to the characteristics of erroneous memories that lead us to believe they are true.

The remainder of this chapter is divided into several sections. The first section provides an overview of memory errors for the content of an episode, specifically looking at work by Loftus and her colleagues on the effects of misinformation. In the second section, research examining errors for the source of memories is reviewed. Johnson and Raye's reality monitoring model of source memory and a review of the research in support of that model is presented in the third section. This is followed by a review of recent work exploring various influences on subjective memory judgments. Much of this work has been done by Jacoby and his colleagues and is discussed in the context of Jacoby's attributional framework. Finally, a discussion of the present research, a summary of theoretical issues, and overview of the experiments is presented.

The "Misinformation Effect": Errors in Memory for Information

An attempt to remember that results in failure can be frustrating. There is, however, an advantage associated with this type of memory error. The absence of information is immediately apparent and steps can be taken to acquire the information through some other means. False memories are a more insidious problem. Strong confidence may be expressed that these memories are the "truth" and unless they are challenged by convincing evidence to the contrary, these sorts of memory errors may never be revealed. Although in many instances the consequences of having inaccurate memories are relatively minor, the problem is more serious when false memories are assumed to be true in eyewitness testimony.

Numerous studies by Loftus and her colleagues (Cole & Loftus, 1979; Loftus, 1975, 1977, 1979; Loftus, Korf, & Schooler, 1989; Loftus, Miller, & Burns, 1978; Wells & Loftus, 1984) and others (e.g., Bekerian & Bowers, 1983; Christiaansen & Ochalek, 1983) have shown how false memories may originate. Using an eyewitness testimony situation, their work demonstrated that new information presented after an event can alter a person's report of the original event. In a typical experiment, subjects viewed an accident videotape that included a stop sign after which they were presented with erroneous information about a yield sign in a questionnaire. Compared with control subjects who received no inconsistent information, subjects who read the inconsistent information were more likely to claim to have seen things that were suggested in that later information. Many variations of the studies have been carried out with a variety of materials. The findings from these experiments support the notion that the process of remembering is a constructive activity. Information from various sources is gathered to form what is called the

Summary of Theoretical Issues and Overview of Experiments

In summary, then, Johnson and Raye's model of reality monitoring depicts source judgments as decisions based on information available in memory from the time of encoding. Characteristic types of information from the encoding episode give rise to particular source decisions. According to this model, source classification errors occur when memories contain characteristics normally associated with another type of source (Finke, Johnson, & Shyi, 1988; Johnson, Raye, Foley, & Foley, 1981; Johnson, Raye, Wang, & Taylor, 1979). In other research, manipulations of current processing have been shown to influence a variety of memory judgments (e.g., Jacoby et al., 1989). Although Johnson et al. have extensively investigated the types of encoding manipulations that affect source decisions, the influence of events that occur after encoding on these decisions has not been systematically studied.

Two sorts of post-encoding influences were investigated in the present research: those present during testing and those that occur between study and test. The effect of a delay on memory performance was also examined, since a delay is typically experienced between the acquisition of information and a task requiring its recollection. Memory for information is often reduced by a delay and poor source memory has been attributed to poor memory in general (Shimamura & Squire, 1991). Others have also reported that source memory accuracy is reduced after a delay (Brown & Halliday, 1991; Schacter et al., 1984; Shimamura & Squire, 1991). In Experiment 1, a preliminary examination the incorrect origin of a particular item of information. The types of errors produced by Loftus's 'misinformation effect' are quite dramatic: subjects claim to have seen objects that were not there. In eyewitness testimony, the repercussions of errors in source memory can be as serious as errors in the content of that memory. In more common, everyday situations, our judgments on the source of a memory may be as important as the content of the memory itself in directing our behaviour. For example, our decision to act or not upon a particular piece of information may be based upon our determination of the credibility of the source who provided that information (Petty & Cacioppo, 1986).

Source memory performance has been an active area of research with memory-impaired subjects (e.g., Schacter, Harbluk, & McLachlan, 1984; Shimamura & Squire, 1987; 1991). Certain pathologies, particularly frontal lobe damage, have been associated with poor source memory performance (Craik, Morris, Morris, & Loewen, 1990; Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Schacter et al., 1984). In a comparison of amnesics and normals, Shimamura and Squire (1991) found that, for amnesic subjects, memory for sources, facts, and events was impaired, but source memory did not appear to be related to either fact or event memory performance. In normals, however, source memory for facts and events depended on the strength of the memory for the facts and events themselves. The particular conditions under which memory for information and its source are tied remain to be worked out and there is still some question whether poor source memory in normal subjects is best characterized as a problem of generally impoverished memory (Johnson, Raye, Foley, & Foley, 1981; Shimamura, 1989).

Age related differences are also found in accuracy for source memory. In general, children and the elderly do not perform as well as young adults on source discrimination tasks (Cohen & Faulkner, 1989; Dywan & Jacoby, 1990; Foley & Johnson, 1985; Hashtroudi, Johnson, & Chrosniak, 1989; Lindsay, 1987; McIntyre & Craik, 1987; Rabinowitz, 1989; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991). These findings are consistent with the idea that the memory disordered, the young, and the elderly are less likely or less able to use metamemory information than healthy young adults. Dywan and Jacoby (1990) suggest that the hesitation of some elderly to act on information may be a result of poor source monitoring. They reasoned that if the origin of information is uncertain, one may be reluctant to act upon that information until the source can be identified and its trustworthiness determined.

"Sources" of information can take many forms. Information may be obtained from printed text, various forms of other media, and other people. One type of source discrimination that is particularly interesting is the distinction between oneself and others as the source of information. When memories result from thoughts, fantasies, dreams, and the products of reasoning and problem solving, we ourselves are the source. Memories associated with the self have a special significance. They are memories unique to us as individuals and are also believed to have special memorial status due to self reference.

When errors are made in the attribution of memories to oneself or others, the results can be quite dramatic. Two physicists attracted media attention to the problem when they made public their disagreement about the events that led to their joint Nobel Prize (Sutton, 1984). The two scientists presented different accounts of who said what during their meetings over several months. Their persistent disagreement about those events led to the dissolution of their collaboration and compelled each of them to write the "true" account of the events to set the record straight.

Bowers and Hilgard (1988) described a case of unconscious plagiarism involving Helen Keller who, as a child, wrote a short story that was published. The story bore a remarkable resemblance to an earlier, published short story. Although she denied copying it and had no recollection of the story telling episode, it was later revealed that the story had been told to her via sign language at an earlier age.

Disagreements about the origins of ideas are also found in claims of copyright infringement that arise in music and the arts (Goldstein, 1989). A composer may honestly believe that a composition is a product of his or her own imagination when in fact the source is a recollection of a copyrighted work. In one famous case, a lawsuit was brought against George Harrison for his song "My Sweet Lord" which bore a strong similarity to the earlier "He's So Fine" by the Chiffons. The judge acknowledged that it was apparent that

Harrison had not been conscious of plagiarizing the theme of the song ("George Harrison Guilty," 1976). The courts, however, are unsympathetic to claims of innocence as a defense to copyright infringement. In a judgment typical of these sorts of cases, one judge ruled that "...he has invaded the author's rights. It is no excuse that in so doing his memory has played him a trick." (Goldstein, 1989, p. 162).

It is even possible not to recognize oneself as the source of an idea. As Skinner (1983) noted "...one of the most disheartening experiences of old age is discovering that a point you have just made- so significant, so beautifully expressed- was made by you in something you published a long time ago" (p.242).

Although most reports of unconscious plagiarism or cryptomnesia (Taylor, 1965) are anecdotal, recently the phenomenon has been examined experimentally (Brown & Murphy, 1989; Brown & Halliday, 1991; Marsh & Bower, 1993). Brown and Murphy (1989), for example, had subjects generate category exemplars, alternating responses with other subjects. Later they were required to recall their own responses along with some new exemplars that had not been generated by anyone within the experimental session. Subjects often plagiarised the responses of others in the group and they were likely to plagiarize from the subjects who preceded them in the generation order. Brown and Murphy suggested that subjects may pay less attention to the response that preceded their own because they are thinking of their own next response. As a result they may lack the information that is required for a correct source attribution. A second possibility is that subjects may be covertly generating answers while they are waiting for their turn (Johnson, Hashtroudi, & Lindsay, 1993). If some of these potential answers are given by others before the subject can produce them, it may be very difficult to distinguish the correct origin of the responses.

Johnson and Raye's Reality Monitoring Model

One of the most comprehensive and successful accounts of source decisions is provided by Johnson and her colleagues (Johnson & Raye, 1981; also Johnson, Hashtroudi, & Lindsay, 1993). Their work focused on the reality status, real or imagined, of past experiences. In the Johnson and Raye model, "real" refers to experiences external to the individual and "imagined" refers to experiences in the individual's imagination. These two categories of experience correspond to different sources of information: others and the self. Johnson and Raye proposed that there are characteristic differences between the memory representations of experienced external events and imagined or self-generated events. One basic distinction concerns the quantity of cognitive operations stored with the memory. For example, information generated by the subject differs from information perceived by the subject in that records of the processes or cognitive operations involved in generating the information remain as part of the memory record of that episode. In comparison, the memory record for information that was perceived contains much less information about cognitive operations because perception is generally a much more automatic process. The representations of internally generated and perceived material differ on other dimensions as well. Perceived information results in memories that are richer in perceptual, sensory, and semantic detail than self-generated information.

Source attributions or origin judgments are decisions made on the basis of information available from the encoding episode (Johnson & Raye, 1981). Memories that are accompanied by semantic, sensory, and perceptual detail with relatively little information about cognitive operations tend to be attributed to an external source. On the other hand, memories that are relatively impoverished in terms of semantic, sensory, and perceptual detail but have an abundance of information about associated cognitive operations, tend to be attributed to internal sources. Confusions or inaccurate classifications of a particular source arise when memories contain characteristics normally associated with another type of source (Finke, Johnson, & Shyi, 1989; Johnson, Raye, Foley, & Foley, 1981; Johnson, Raye, Wang, & Taylor, 1979; see also Durso & Johnson, 1980). The decision process may be influenced by the metamemory assumptions or biases that people hold about how memory works. Johnson and Raye note, for example, that subjects often believe that self-generated information is more memorable and that they would remember if an idea had been theirs. Factors such as the amount of time available, availability of relevant information in memory, and the cost of mistakes may also influence the source decision process.

Much of the previous research on reality monitoring has relied on manipulations of cognitive operations. In these experiments, subjects must decide whether a particular word was one they read or one they generated. For example, subjects were more accurate in later identification of origin when they were required to generate category instances without first letter cues than if those cues were provided (Johnson, Raye, Foley, & Foley, 1981). According to Johnson et al. (1981), the relative ease of generating information to a cue results in fewer cognitive operations which in turn provide less information on which to base source discriminations. In their second experiment, Johnson et al. (1981) found that the correct identification of origin was increased when the study task required a less typical exemplar of a category compared with generations that required a more common exemplar of a category. In both of these experiments, source identification was more accurate in the conditions where greater cognitive effort was required to generate the item (generating without a first letter clue or generating a nontypical exemplar).

Repetition of certain types of processing can also influence source accuracy, but the specific effects depend on the particular conditions involved. Generating an item twice at encoding, for example, increases the correct identification of origin (Johnson, Raye, & Durso, 1980) presumably because an increase in cognitive operations is likely to result from repeated generation.

Repeated presentation of perceived items, however, increases source confusions (e.g., Johnson, Raye, Wang, & Taylor, 1979; Johnson, Taylor, & Raye, 1977). The second presentation provides an opportunity for increased elaboration and an internal source attribution becomes more likely when there is more self-generated information associated with the memory.

According to reality monitoring theory (Johnson & Raye, 1981; also Lindsay, 1987), source discriminations should be easier to make between two memories with different types of sources than between two memories with internal sources or two memories with external sources. This prediction is based on the idea that the operations involved with two memories from the same type of source (internal or external) would be more similar than two memories from different types of sources. Foley, Johnson, and Raye (1983) found that when subjects engaged in two forms of self-generated activity (saying an item aloud or thinking about it) they were less accurate in deciding if they had said it or only thought about it than if they had to discriminate between words they or someone else had said. Raye and Johnson (1980) found that subjects had more difficulty identifying the origins of memories of two perceptions than identifying the origins of memories of saying or hearing a word. Lindsay (1987; Lindsay & Johnson, 1991) has shown that increased source attribution errors are likely if the two sources are similar on other dimensions such as medium of presentation, perceptual properties, or semantic content. Anderson (1984) reported similar findings for confusions for memories of doing an activity and memories of imagining doing an activity, such as tracing an outline of an object.

In summary, reality monitoring theory provides an account of how decisions about the origin of memories are made. At encoding, different sources (the self or others) produce characteristic types of information. Later, when a source discrimination is required, this information from encoding is assessed and a source decision is made. According to this theory, source decision errors arise when a memory is accompanied by information that is usually associated with the other type of source.

<u>The Influence of Testing Conditions on Memory Judgments</u>

Johnson and Raye manipulated encoding conditions in their investigation of source memory, but some recent work shows that the conditions at testing may also influence several types of memory judgments. These findings suggest that testing conditions might affect source memory as well. Much of the work in this area has been done by Jacoby and his colleagues, who demonstrated the influence of processing manipulations on several types of subjective memory judgments (for a review see Jacoby, Kelley, & Dywan, 1989).

Jacoby (Jacoby et al., 1989) has provided an attributional framework for understanding the experience of remembering. He emphasizes the role of subjective experience in remembering and argues that the process of remembering is different from retrieving a copy of the original event. In this framework, remembering is the act of attributing something to our past. There may not necessarily be a correspondence between subjective experience and a representation in memory. In support of this view, he describes situations where mismatches occur between the objective record and subjective experience. Amnesics, for example, may not experience remembering but are able to act upon recently acquired information by completing word stems (Warrington & Weiskrantz, 1974) or answering questions (Schacter et al., 1984). A second line of evidence comes from Jacoby's own research demonstrating illusions of memory in subjects with normal memory.

Jacoby and his colleagues have examined the role of subjective experience in remembering in a series of studies. Their general strategy was to try to influence these judgments by manipulating the ease of perceptual processing at test. This choice of manipulation is based on the logic that, if an item is processed more easily than other items, subjects may use this relative fluency to infer past experiences (see Begg, Duft, Lalonde, Melnick, & Sanvito, 1989, for the influence of ease of processing on memory predictions).

Perceptual fluency has been shown to influence memory judgments of a variety of characteristics including fame, familiarity, loudness, brightness, and prior presentation (e.g., Jacoby, Allan, Collins, & Larwill, 1988; Jacoby & Whitehouse, 1989; Jacoby, Woloshyn, & Kelley, 1989; Wittlesea, Jacoby, & Girard, 1990). Wittlesea, Jacoby, and Girard (1990), for example, investigated several illusions of memory created by perceptual ease. In one experiment, subjects were more likely to claim that a word had been repeated if it was presented in light rather than heavy visual noise. In another experiment, subjects judged a word to be more perceptually clear if it had been viewed earlier, although no actual manipulation of visual clarity was made.

Jacoby and Whitehouse (1989) used manipulations of perceptual fluency to induce "recognition" of new words. During a test of recognition memory, context words preceded most of the items. When the context word was identical to a subsequent "new" word and was presented too quickly to be identified, subjects were influenced to "recognize" that new word as old. It is interesting to note that when subjects were aware of the context words, due to longer presentation durations, they were not susceptible to this false recognition effect. Jacoby and Whitehouse discussed their findings in terms of how attribution processes underlie the feeling of familiarity. In their view, subjects were able to discount the feeling of familiarity of the test word when they could attribute its familiarity to the recent presentation of the context word.

Kelley, Jacoby, and Hollingshead (1989) presented subjects with words to hear or read. Later they were tested for identification of these words at brief durations. They found that regardless of its actual status as "read," "heard," or "new," words that were perceptually identified during brief presentations were more likely to be called "read" than words that were not identified. Kelley et al. argued that the basis for these decisions was the subjective feeling of relative fluency for those words that were perceptually identified.

Lindsay and Kelley (1991) varied the relative ease of processing with a different type of task. Subjects studied a long list of words and were later tested with word fragments that were easy to complete (1 missing letter) or difficult to complete (2 missing letters). Subjects generated the completions and then rated the words as "clear memory", "feels familiar", or "no memory". There were relatively more reports of "clear memory" and "feels familiar" with the easy-to-complete fragments, and the effect occurred for both hits and false alarms.

Across these various types of decisions, the manipulations had similar effects: the way in which the information was presented at the time of decision (which determined how the information was to be processed) influenced the actual decision about the earlier experience with the word.

In the above studies, perceptual fluency was manipulated during item presentation so as to influence subjective memory judgments. Some recent work by Rabinowitz (1990) indicates that generating words at retrieval can also influence subjects' judgments about the original presentation format of those words. Subjects read or generated words from fragments (presented with semantically related cues) at both study and test. At the test, subjects were required to indicate whether the word had been presented for study (a word recognition judgment) and, if they recognized the word, whether it had been initially read or generated (an origin judgment). Rabinowitz hypothesized that generating the word again at retrieval would facilitate identification of similar procedures that took place during encoding, enabling the subjects to make accurate origin decisions. Subjects would recognize that they had "done this before" when generating at test and so make accurate source attributions. For items that had been originally read at study, no particular benefit was expected from reading again at test (because reading is a relatively automatic process), but when subjects generated at test, origin judgments for items that were originally read were expected to be accurate because generating these items at test would result in a mismatch with the information from encoding.

The important finding for present purposes is that subjects' judgments of origin were influenced by the processing of the information at the time of testing. Contrary to Rabinowitz's prediction that generating at test would lead to accurate source decisions for both originally read and generated words, when subjects generated originally read words at retrieval, they often erred and said the word was originally generated. Subjects did not appear to compare and contrast information from encoding and retrieval to guide their source decisions, but apparently found it difficult to distinguish between the two sets of operations and made errors about the original format of the words.

In summary, the studies reviewed in this section provide examples the ways in which testing conditions can influence memory judgments and highlight the importance of investigating the influences that later experiences as well as study conditions can have on memory judgments.

The Present Research

The type of source memory investigated in this thesis involves the discrimination between oneself and others as the source of information. These experiments were motivated by both applied and theoretical considerations. The questions raised by applied memory researchers investigating memory for the content of an episode are also relevant for the study of source memory. In an attempt to determine whether or not memory reports were trustworthy, much of the previous research was directed toward assessing the accuracy of memory performance under various conditions. The evidence from the eyewitness research indicated that subjects' reports for an event were not necessarily reliable. The empirical evaluation of memory accuracy was particularly important because these reports were commonly believed to be accurate. The main focus of the present research is the accuracy of judgments for the origins of information in memory and the conditions under which those reports are more or less likely to be accurate.

A second question of interest, also raised in the applied memory literature, concerns the relationship between the accuracy of reported memories and subjects' confidence in those reports. This connection underlies two central issues in the eyewitness literature (Orne, Whitehouse, Ginges, & Orne, 1988). The first issue involves the relationship between the displayed level of confidence and the accuracy of the information provided by the witness. The second issue concerns the influence on juror perceptions of the level of confidence that the witness conveys. Previous research has demonstrated that witness confidence may bear little relationship to the accuracy of the content of eyewitness information (Wells, Lindsay, & Ferguson, 1979). The problem is compounded by additional findings indicating that the level of confidence exhibited by the witness is a major determinant of perceived eyewitness credibility (Lindsay, Wells, & Rumpel, 1981; Wells, Lindsay, & Ferguson, 1979).

In the present research, the relationship between source attribution accuracy and subject confidence in those decisions was investigated. The confidence one has in a decision contributes to the decision to act (or not) on that information (Jacoby, Kelley, & Dywan, 1989). Thus, confidence provides the basis for action as well as influencing the perceptions of others. If the confidence subjects display for source attributions reflects the accuracy of those decisions, then confidence could provide a useful tool in the evaluation of source decisions. If, however, the relationship between confidence and accuracy is unreliable, as is sometimes the case with eyewitness reports, then these reports should be interpreted with caution.

The fundamental issue from a theoretical perspective is how to best characterize the source decision process. Johnson and Raye (1981) proposed that this particular type of source decision cannot be easily accommodated within traditional models of memory and they argued that a separate theory, reality monitoring, was required to accommodate their research findings. In their view, one does not determine the origin of a bit of information by reading off a label or associated tag. Rather, Johnson and Raye depict the source decision process as a reconstructive one where subjects examine information from encoding and make source decisions on the basis of that information.

According to Johnson et al. (1981) the source decision process in reality monitoring operates on information preserved in memory over a substantial retention interval. Yet there is persuasive evidence from other research that source memory decisions, like other sorts of memory judgments, may be influenced by additional factors. Jacoby et al. (1989) argue that many types of memory judgments are influenced by information that is not part of the memory representation itself. Rather, the characteristics of our later experience may influence our judgments of earlier experience. Loftus (1991b) has also argued that due to reconstructive processes in memory our "memories" of past experiences are often an amalgam of information gathered from a variety of sources.

The purpose of the present series of experiments was to systematically examine the effect that events that occur after encoding might have on source memory attributions. While subjects may typically rely on the information from the encoding episode for their source decisions (Johnson et al., 1981), Rabinowitz (1990) and Jacoby (Jacoby et al., 1988; Jacoby & Whitehouse, 1989; Jacoby, Woloshyn, & Kelley, 1989; Wittlesea et al., 1990) have shown that the conditions under which memory is tested can influence several other types of memory judgments. These finding suggest that post encoding factors may also influence the accuracy of source attributions. Specifically, it was predicted that source memory confusions would arise when the type of processing required of the information during testing, or the period between study and testing, is representative of another type of source. If this is the case, then a characteristic pattern of source errors should be produced. Source accuracy for items generated at study should be reduced when these words are tested under conditions representative of an external source (word identification). In a similar fashion, source accuracy for words read at study should be reduced when testing conditions produce information that mimics internal sources.

Summary of Theoretical Issues and Overview of Experiments

In summary, then, Johnson and Raye's model of reality monitoring depicts source judgments as decisions based on information available in memory from the time of encoding. Characteristic types of information from the encoding episode give rise to particular source decisions. According to this model, source classification errors occur when memories contain characteristics normally associated with another type of source (Finke, Johnson, & Shyi, 1988; Johnson, Raye, Foley, & Foley, 1981; Johnson, Raye, Wang, & Taylor, 1979). In other research, manipulations of current processing have been shown to influence a variety of memory judgments (e.g., Jacoby et al., 1989). Although Johnson et al. have extensively investigated the types of encoding manipulations that affect source decisions, the influence of events that occur after encoding on these decisions has not been systematically studied.

Two sorts of post-encoding influences were investigated in the present research: those present during testing and those that occur between study and test. The effect of a delay on memory performance was also examined, since a delay is typically experienced between the acquisition of information and a task requiring its recollection. Memory for information is often reduced by a delay and poor source memory has been attributed to poor memory in general (Shimamura & Squire, 1991). Others have also reported that source memory accuracy is reduced after a delay (Brown & Halliday, 1991; Schacter et al., 1984; Shimamura & Squire, 1991). In Experiment 1, a preliminary examination of testing influences on the accuracy of source memory was undertaken. To anticipate, source decision accuracy was reduced when words that were generated at study were tested using word identification methods. In Experiment 2, the conditions at test were manipulated within subjects to examine their influence on word recognition, decisions about the source of that information, and ratings of confidence made for those decisions. Here, source attributions were less accurate when words read at study were presented for testing in a generation format that was indicative of the alternative source. Testing conditions also exerted an influence on confidence ratings but only for incorrect source decisions. The primary purpose of Experiment 3 was to replicate these findings and to assess the relationship between subjects' confidence for recognition memory and the accuracy of their source decisions. Stronger memories have been shown to be less vulnerable to distortion, at least under certain conditions (e.g., Shaughnessy & Mand, 1982). As in Experiment 1, source decision accuracy was reduced for words generated at study and tested using word identification. Also the relationship between rated recognition confidence did not prove to be a reliable predictor of source decision accuracy. Finally, it has been well documented that postevent information can reduce the accuracy of reports of earlier events (e.g., Loftus, 1991b). In Experiment 4, inconsistent information that occurs between study and test was found to reduce the accuracy of source decisions for words that had been generated at study.

Taken together the results of these experiments should provide more information about the accuracy of source memory, the reliability of confidence for source decisions, and add to our knowledge of how source attributions are made.

CHAPTER 2

Experiment 1:

Preliminary Examination of Testing Influences on Source Memory Accuracy

Source judgment errors may underlie many everyday memory problems and are also implicated in mistaken eyewitness testimony and cases of unconscious plagiarism where a distinction must be made between oneself and others as the source of information (Johnson, Hashtroudi, & Lindsay, 1993). One cause of these sorts of errors is when memories contain information from encoding that is usually associated with the alternative source (Johnson & Raye, 1981; Johnson, Hashtroudi, & Lindsay, 1993). In several studies Johnson and her colleagues (e.g., Finke, Johnson, & Shyi, 1988; Johnson, Raye, Foley, & Foley, 1981) demonstrated that subjects were more likely to attribute a memory to an internal source, for example, if that memory was associated with cognitive effort. The focus of the present study was the role that later influences, specifically the conditions at testing, play in decisions of origin. Manipulations of relative perceptual fluency at the time of testing, for example, are known to influence several other types of memory judgments (Jacoby, Allan, Collins, & Larwill, 1988; Jacoby & Whitehouse, 1989; Kelley, Jacoby, & Hollingshead, 1989).

Following previous work, an attempt was made in the present experiment to use two of the dimensions that Johnson and Raye (1981) indicated were important for source judgments. Words were read and generated at study to provide memories with external and internal sources. Later, subjects were tested for their memory of the information and its source in one of three test conditions: word identification, generation, or recognition. According to the Johnson and Raye model, cognitive operations are usually indicative of the self as the source and perceptual clarity (as well as lack of cognitive operations) is usually taken as evidence that a memory had an external source. The first two testing conditions were specifically chosen to provide processing information representative of specific types of sources. In the generation testing condition, subjects performed the same cognitive operations that at study had provided evidence for the self as the source. In contrast, reading a word in the word identification condition results in a very fluent perception, the type of information consistent with an external source. The recognition testing condition was included to provide a comparison with the other two testing formats. Although it is difficult to determine what an appropriate control or neutral condition might be for this sort of task, recognition differs from the other two test conditions in that it lacks the specific characteristics on which the two type of source judgments are based.

In summary, subjects studied words by reading or generating them to produce memories representative of external or internal sources. Memory for these words and their sources was tested under conditions which were consistent with the study presentation, representative of the alternative study presentation, or under recognition conditions. It was predicted that source memory accuracy would be reduced when memory was tested under conditions that were representative of the other source.

Method

Subjects and Design

Ninety University of Western Ontario undergraduates participated for course credit or payment of $5.00. A 2 X 3 factorial design was used where study condition (read and generate) was a within-subjects factor and testing condition (word identification, generation, recognition) was a between-subjects factor.

Materials

A group of 114 nouns (6 or 7 letters in length), of moderate frequency (mean frequency = 35.17 occurrences per million), was obtained from Kučera and Francis (1967). Twenty-four of these words were randomly selected for use in the practice procedure for the word identification task described below. The remaining 90 words were randomly allocated to three lists of 30 words, each of which served equally often in each of the two study conditions and as distracters at test. The complete list of materials is presented in Appendix A.

Procedure

There were two main components to the procedure, study and testing. The conditions for studying the materials were identical for all subjects, but subjects participated in only one of the testing conditions. All subjects were tested individually.

Study phase. Subjects were presented with words to read (e.g., HORIZON) or to generate from an anagram (e.g., OHRIZON) that required the reversal of the first two letters to solve. The 60 words (half read, half generated) were presented in capital letters on a computer monitor for 3 s each with 1 s ISI. No more than three read or generate items appeared consecutively. Subjects were instructed to read the words aloud. After the study list was completed, subjects performed a state generation task for 5 min where they were required to write the state names of the USA.

Test phase. Subjects participated in one of three testing conditions: word identification, generation, or recognition. A description of each test condition is presented below. A single list order was used for testing where no more than three occurrences of read, generate, or new words appeared consecutively. For each item, regardless of test presentation format, the subject read the word aloud and then made a recognition judgment. For recognized items, the subject made a further decision as to whether that word was read or generated at study. Subjects worked at their own pace and responses were recorded by the experimenter.

Word identification condition. Word identification items were presented on the monitor in their normal format (e.g., HORIZON). The sequence of events for one Word Identification trial was as follows. A keypress initiated the trial and 2 s later a fixation cross (+) was displayed for 500 ms in the middle of the monitor. A pattern mask consisting of eight ampersands (&&&&&&&&) was displayed for 150 ms. The word was then displayed for the designated time, followed by a second pattern mask for 150 ms.

The appropriate display duration was individually determined for each subject in this condition prior to the study phase of the experiment. Subjects were presented with 24 practice words and the display duration was adjusted until the subjects could identify approximately 40% of the items. This procedure resulted in display durations of either 16 ms or 33 ms (mean presentation duration was 23.93 ms).

Generate condition. All words tested in this condition were presented with the first two letters reversed (e.g., OHRIZON), the same format as the generated items at study. The words for generation test were presented on a printed page where the subjects worked with a cardboard mask that allowed only one item to be viewed at a time.

Recognition condition. The words for recognition test were presented on a printed page in normal format (e.g., HORIZON). Subjects used a cardboard mask that allowed only one word to be viewed at a time.

Instructions for the word identification practice, and all three testing conditions are presented in Appendix B.

Results

The Bonferroni procedure was used for nonpredicted tests of means and predicted comparisons of means were evaluated at an alpha level of .05 (Kirk, 1968).

Recognition Memory

The use of the word identification task required that each word be identified (read aloud) before a recognition judgment could be made. As found in previous work (e.g., Jacoby & Dallas, 1981), words that had been studied were identified more often than new words. Comparable levels of identification were observed for words from both study conditions (read=.58, generate=.56; $t_{(59)}<1.00$, $p>.05$). Fewer of the new words (.40) were identified than previously studied words ($ts \geq 7.04$, $p<.0001$).

The mean proportion of words recognized in each of the various study and test conditions is presented in Table 2.1. Investigation of the Study X Test Condition interaction ($F_{(2,87)}=5.92$, $p<.003$, $MSe=.00753$) indicated that testing conditions affected recognition for words read at study but not those generated at study. Words read at study were not recognized as well in the generate test condition (.68) as in the word identification (.78) and recognition (.74) test

Table 2.1

Word Recognition as a Function of Study Condition and Test Condition

[Experiment 1]

Study Condition	Test Condition		
	Word Identification	Generation	Recognition
Read	.78	.68	.74
Generate	.88	.88	.88

conditions ($ts \geq 2.68$, $p < .005$). Words read at study were better recognized in the word identification condition than the recognition condition, although this difference was significant at the $p = .05$ level but not by the Bonferroni procedure ($t_{(87)} = 1.79$, $p < .05$). The lack of effect for generated words may be due to a ceiling effect (all means = .88; all $ts < 1.00$). A generation effect was displayed in a main effect for study condition where generated words (.88) were recognized better than read words (.73; $F_{(1,87)} = 129.83$, $p < .0001$, $MSe = .00753$).

Source Memory

Table 2.2 presents the source memory data, which are based on the correct source attributions for recognized words. Consistent with previous research (e.g., Johnson & Raye, 1981) subjects made more correct source attributions for words read at study (.80) than for words generated at study (.66; $F_{(1,87)} = 41.95$, $p < .0001$, $MSe = .02201$). The nature of the source decision process and subjects' metamemory assumptions contribute to better source decision accuracy for externally presented information than for internally generated memories (Johnson & Raye, 1981). Subjects base their source decisions for internally generated information on an assessment of the cognitive operations that accompany the memory. If the cognitive operations associated with a memory are relatively few compared with other memories or if there is uncertainty associated with the decision, then subjects tend to attribute the memory to an external source by default. Johnson and Raye (1981) characterized the decision process as "I would know if it had been my own

Table 2.2

Source Accuracy as a Function of Study Condition and Test Condition

[Experiment 1]

Study Condition	Test Condition		
	Word Identification	Generation	Recognition
Read	.80	.80	.80
Generate	.60	.68	.68

idea." Memories originating from an external source, however, are attributed to an external source, regardless of the subject's level of certainty. Neither the Study X Test interaction ($F_{(2,87)}$=1.15, p>.05, MSe= .02201) nor the main effect for Test condition ($F_{(2,87)}$=1.66, p>.05, MSe=.02011) were significant. However, because a difference in source accuracy due to testing conditions was predicted, differences among means were further examined. For words read at study, no differences in source attribution accuracy were observed for words read at study (all means =.80), possibly reflecting a ceiling effect. Subjects, however, were less likely to correctly identify the source of generated items when those words were tested by word identification than either of the two other testing formats (ts=2.09, p<.05). None of the other comparisons of testing conditions for words studied by reading or generating approached significance (ts<1.00).

Source Bias Analyses

Another way to investigate the effects of testing conditions on subjects' decisions about the origins of information is to examine their response bias using the methods of signal detection theory (Macmillan & Creelman, 1991). Following Rabinowitz (1990), the proportion of read judgments for items that were initially read [p("R"|R)] was treated as a hit rate and the proportion of read judgments for items that were initially generated [p("R"|G)] was treated as a false alarm rate. For each subject in each testing condition, two values were calculated from the hit and false alarm rate: d' and C (a measure of response bias). Snodgrass and Corwin (1988) recommend the use of C as a measure of bias because they found d' and C to be independent whereas d' and beta are not. Negative C values indicate a bias towards answering "read" and positive C values indicate a bias towards answering "generate".

Measures of sensitivity and bias for source decisions are presented in Table 2.3. Sensitivity measures (d') ranged from 1.42 to 1.64 and did not vary significantly across testing conditions ($F_{(2,87)}$=.44, p>.05, MSe=.89006). As predicted, the values for C showed the greatest bias towards read responses in the word identification (-.43), followed by generation (-.33), and recognition (-.19) conditions. Although none of the differences were significant by ANOVA ($F_{(2,87)}$=1.90, p=.15, MSe=.23399), predicted t tests indicated a greater bias toward read responses in the word identification condition compared to the recognition condition ($t_{(87)}$=1.94, p<.05, other ts≤1.13 p>.05). This finding is consistent with the data for source decisions where source accuracy was reduced when generated words were tested under conditions of word identification.

New Item Analyses

Testing conditions did not influence subjects' ability to identify unstudied words (means=.81-.82; $F_{(2,87)}$=.05, p>.05, MSe=.00115). When new words were mistakenly identified as old, subjects tended to attribute these words to read sources but no significant differences among testing conditions were observed (means=.77-.85; $F_{(2,77)}$=.72, p>.05, MSe=.05411).

Table 2.3

Sensitivity and Bias Measures for Test Conditions

[Experiment 1]

Measure	Test Condition		
	Word Identification	Generation	Recognition
d'	1.47	1.64	1.42
C	-.43	-.33	-.19

Discussion

The results of Experiment 1 provide preliminary evidence that the conditions of memory testing may influence memory judgments. Recognition accuracy was reduced when words read at study required generation at test. Lindsay and Kelley (1991) report similar findings for word recognition. Source accuracy was reduced for words generated at study when they were tested under conditions of word identification. In addition, subjects tested by word identification showed greater bias to say that an originally generated word had been read than subjects who were tested by a recognition test. These findings provide support for the idea that subjects may be influenced by the conditions of testing when making source memory judgments. Similar influences have been observed with other types of memory judgments (e.g., Jacoby et al., 1988; Jacoby & Whitehouse, 1989; Kelley, Jacoby, & Hollingshead, 1989).

Nevertheless, testing conditions did not always show the expected influence on source decisions. In particular, the accuracy of source decisions for words read at study was not reduced when testing conditions required generation. In previous research, Rabinowitz (1990) found that subjects were biased towards generate judgments when materials were generated at test. In the present study source decision accuracy for words that had been read was equally accurate across all testing conditions. Although the possibility of a ceiling effect cannot be ruled out, an alternative interpretation is that the testing conditions themselves may have imposed a constraint on source decisions.

Subjects were aware that all of the words at test were not studied in the format in which they were presented at testing. As a result, they may have discounted much of the influence of the testing condition and relied on the information from encoding for their source decisions. For this reason, between subject testing, where subjects are exposed to only one type of testing influence, may have underestimated the influence of testing conditions on source decisions. It could also be argued that a situation where memories come back more or less quickly and with more or less effort is more similar to what we experience in everyday memory. These possibilities were examined in Experiment 2, where all materials were presented under all testing conditions to all subjects.

CHAPTER 3

Experiment 2:

The Influence of Testing Conditions on

Memory for the Source of Information and

Confidence in those Judgments

The results of Experiment 1 demonstrated that judgments about the origin of a recognized word could be influenced by the way that word was processed at the time of the judgment. Specifically, subjects were less likely to make a correct source judgment for a generated word when the test format was word identification. That subjects were more likely to call these words "read" is consistent with the idea that they were relying on information from the testing rather than encoding episode for the source decision. Word identification provides a fluent perceptual experience, the type of information that is usually characteristic of an external rather than internal source.

The other testing conditions, however, failed to produce any change in source decisions. Most importantly, the generation test condition did not reduce source decision accuracy for words that had been read at study as would have been expected from previous research (e.g., Rabinowitz, 1990). One important difference between Experiment 1 and Rabinowitz (1990) was that subjects in Experiment 1 experienced only one of testing formats whereas subjects in Rabinowitz (1990) experienced both test formats.

As in Experiment 1, subjects in the present study read or generated words at study. Later, they were tested for their memory of the information and its source in three different conditions: word identification, the same generation task as at study, and a new more difficult generation task. The important difference in the present experiment was that subjects experienced all three types of test format. The testing conditions were chosen to provide processing information representative of specific types of sources. Reading a word in the word identification condition results in a very fluent perception, the type of information consistent with an external source. In the generation-same test condition, subjects performed the same cognitive operations that at study had provided evidence for the self as the source. The generation-difficult test condition was included so that the influence of a different type of processing at the time of test could be observed and was expected to produce a level of source performance between the other two test conditions given that it was not directly associated with either of the two sources from study. The effect of processing difficulty on word recognition was investigated by Peynircioğlu and Watkins (1988; Watkins & Peynircioğlu, 1990; Peynircioğlu & Tekcan, 1993) who showed that there may be a bias to say an item is old if that item is obscured or disguised and has to be revealed before a recognition judgment can be made. In contrast, other research by Lindsay and Kelley (1991) has shown that a word is more likely to be judged old when it is processed with relative ease. Given these contradictory findings, the exact influence of this manipulation on word recognition was unclear.

In summary, subjects studied words by reading or generating them to produce memories representative of external or internal sources. They were tested for memory of these words under conditions which were consistent with the study presentation, representative of the alternative study presentation, or in a completely new format. As in the previous experiment, it was predicted that source accuracy would be reduced when subjects were tested in a format that was representative of the alternative type of source. The level of source accuracy performance for the third test condition was expected to fall between the other two conditions.

The confidence displayed by subjects for their source judgments was an additional focus of interest. Confidence judgments have been shown to be poor indicators of accuracy in eyewitness research (Wells et al., 1979). In the present study, confidence ratings were collected for subjects' source decisions under the various conditions of testing to determine the relation between source decisions and subjective ratings of confidence for those decisions.

Finally, memory was tested immediately after study or after a 24-hour delay. In general, the passage of time reduces memory accuracy and eyewitness memory errors due to the misinformation effect are more likely to occur after a delay (Loftus, Miller, & Burns, 1978). Testing after a delay is known to reduce accuracy of other forms of source memory as well (Brown & Halliday, 1991; Schacter et al., 1984; Shimamura & Squire, 1991).

Method

Subjects and Design

Thirty-six University of Western Ontario undergraduates participated for course credit. A 2 X 2 X 3 factorial design was used with study conditions (read and generate) and testing conditions (word identification, generation-same, generation-difficult) manipulated as within subjects factors and the time of test (immediate or 24-hour delay) manipulated as a between subjects factor.

Materials

Additional words selected from Kučera and Francis (1967) were added to the materials from Experiment 1 to form a pool of 135 words. The words were of moderate frequency (mean frequency = 35.25 occurrences per million) and were 6 or 7 letters in length. The words were randomly allocated to three lists of 45 words, each of which served equally often in each of the two study conditions and as distracters at test. The complete list of materials is presented in Appendix C.

Procedure

Subjects were tested individually. There were two main components to the procedure, study and testing. All subjects studied words in both conditions and were tested using the three testing formats, described below.

Study phase. Subjects were presented with words to read (e.g., HORIZON) or generate from an anagram (e.g., OHRIZON) that required the reversal of the first two letters to solve. The 90 words (half read, half generated) were presented in capital letters on a computer monitor for 3 s each with 1 s ISI. No more than three read or generate items appeared consecutively. Subjects were instructed to read the words aloud.

Test phase. Prior to testing, the display duration for the word identification test was individually determined for each subject using the procedure and materials described in Experiment 1. Testing occurred immediately following study or after a 24 hour delay. Within each of the three testing formats, 1/3 of the words were from each of the study conditions and 1/3 of the words were new. A single list order was used for testing, but the testing condition orders were counterbalanced. Items were presented for testing in groupings of 7 or 8 items of each test type. Word identification items were presented on the monitor and words for the two generation tasks were presented on a printed page where the subjects worked with a cardboard mask that allowed only one item to be viewed at a time.

For each item, regardless of test presentation format, the subject read the word aloud and then made a recognition judgment. For recognized items, the subject made a further decision as to whether they had read or generated the word at study and rated their confidence for that judgment using a 1 (guess) to 5 (certain) scale. Their responses were recorded by the experimenter.

Word identification condition. The sequence of events for one word identification trial was as follows. A keypress initiated the trial and 2 s later a fixation cross (plus sign) was displayed for 500 ms in the middle of the monitor. A pattern mask consisting of eight ampersands (&&&&&&&&) was displayed for 150 ms. The word was then displayed for the designated time, followed by a second pattern mask for 150 ms. All words presented for word identification were in their normal format (e.g., HORIZON).

Generation-same condition. Items tested in this condition were presented with the first two letters reversed (e.g., OHRIZON), the same format as the generated items at study.

Generation-difficult condition. In this condition, subjects were required to generate the items from complicated anagrams (e.g., ZIONROH). The anagrams were formed by assigning the numbers 1-7 to the letters in the word (or 1 to 6 for 6 letter words) and then reordering the letters in the order 5467321. A card displaying the rule was available for subjects to consult.

The full set of instructions is presented in Appendix D.

Results

Pooled error terms and Satterthwaite's adjustment for degrees of freedom were used in tests of simple main effects from interactions involving repeated measures (Winer, 1962). Predicted means were evaluated at an alpha level of .05 and the Bonferroni procedure was used for nonpredicted tests of means (Kirk, 1968).

Recognition Memory

Subjects attempted to read each word aloud in the word identification task before making a recognition judgment. As found in Experiment 1 and as expected from previous research (e.g., Jacoby & Dallas, 1981), words that had been studied were identified more often than unstudied words. Comparable levels of identification were observed for both of the study conditions (read =.76, generate =.73; $t_{(68)}=1.18$, $p>.05$). Fewer of the new words (.67) were identified than words from either of the study conditions ($ts \geq 2.35$, $p<.05$) ($F_{(2,68)}=5.74$, $p<.005$, $MSe=.01173$).

The recognition data are summarized in Table 3.1. A 2 (read or generate) X 3 (word identification, generation-same, generation-difficult) X 2 (immediate or delayed testing) ANOVA was used to analyze the data. The typical generation effect (Slamecka & Graf, 1978) was demonstrated in that items generated at study (.84) were recognized more often than items read at study (.73) ($F_{(1,34)}=37.69$, $p<.0001$, $MSe=.01728$). A significant main effect for test condition indicated that words presented in the generation-difficult testing format (.76) were not recognized as well as those presented in the word identification (.80) or the generation-same (.79) format ($F_{(2,68)}=3.24$, $p<.05$,

Table 3.1

Word Recognition as a Function of Study Condition and Test Condition

[Experiment 2]

Study Condition	Test Condition		
	Word Identification	Generation Same	Generation Difficult
Read	.75	.73	.70
Generate	.86	.84	.81

MSe=.01262). Subjects were more likely to recognize studied words under the word identification than the generation-difficult condition This comparison was significant at the $p=.05$ level but did not hold using the more stringent Bonferroni procedure ($t_{(68)}=1.78$, $p<.05$). Neither of the other two pairwise comparisons differed ($t_{(68)}\leq1.21$, $p>.05$). The 24 hour delay in testing produced no reduction in recognition memory (immediate test = .78, delayed test = .79; $F_{(1,34)}= .11$, $p>.05$, MSe=.05322).

Source Memory

Table 3.2 presents the source memory data, which are based on the correct source attributions for items recognized. These data were analyzed in an ANOVA with the same factors as in the word recognition analysis.

The most important finding for the source accuracy data was a significant Test X Study interaction ($F_{(2,68)}=7.37$, $p<.001$, MSe=.03546). When items were generated at study, source attributions were most accurate for items tested in the generation-same format (.60) and this differed significantly from both the generate-difficult (.53) and the word identification (.51) conditions ($ts\geq1.84$, $p<.05$). The generation-difficult and word identification conditions did not differ ($t_{(130)}=.35$, $p>.05$). A different pattern emerged for items originally read. Source attributions were slightly more accurate for items tested in the word identification (.78) than the generation-difficult (.74) conditions, but not significantly so ($t_{(130)}=1.07$, $p>.05$). Source attributions were least accurate in the generation-same format (.64), which differed from both the word

Table 3.2

Source Accuracy as a Function of Study Condition and Test Condition

[Experiment 2]

Study Condition	Test Condition		
	Word Identification	Generation Same	Generation Difficult
Read	.78	.64	.74
Generate	.51	.60	.53

identification and generation-difficult conditions ($ts \geq 2.49$, $p < .01$).

Two other effects were significant. Delayed testing reduced source accuracy (immediate = .67, delayed = .60; $F_{(1,34)}=7.41$, $p<.01$, MSe=.04085). Overall, study condition influenced source accuracy such that source decisions were more accurate for items read (.72) than for items generated (.55) at study ($F_{(1,34)}=19.42$, $p<.0001$, MSe=.08354). This finding replicates similar results reported by Johnson and her colleagues (e.g., Johnson, Raye, Foley, & Foley, 1981).

Source Bias Analyses

The source bias analyses are presented in Table 3.3. The sensitivity measure (d') was not influenced by the testing conditions ($F_{(2,68)}=1.00$, $p=.37$, MSe=.60905) although sensitivity was reduced after a 24 hour delay (immediate = 1.16, delay =.62; $F_{(1,34)}=5.65$, $p<.02$, MSe=1.39854). The interaction of testing condition and time of test was not significant ($F_{(2,68)}=.60$, $p=.54$, MSe=.60905).

Testing conditions influenced subjects' bias to respond "read" for the source of words they recognized ($F_{(2,68)}=7.62$, $p<.01$, MSe=.27494). Negative values of C indicate a bias towards answering "read" and positive C values indicate a bias towards answering "generate". The bias to report a word as read was strongest for words tested by word identification (-.53), weakest for words tested by generation-same (-.06), and the generation-difficult condition resulted in bias measure between these two values (-.38). Predicted t tests

Table 3.3

Sensitivity and Bias Measures for Test Conditions

[Experiment 2]

	Test Condition		
Measure	Word Identification	Generation Same	Generation Difficult
d'	1.00	.74	.93
C	-.53	-.06	-.38

confirmed this pattern of bias ($ts \geq 2.59$, $p < .01$) but the comparison for word identification and generation-difficult failed to reach significance ($t_{(68)} = 1.22$, $p = .11$). Neither the main effect for delay nor the interaction of delay and test were significant ($Fs < 1.00$, $p > .05$).

Confidence Ratings for Source Attributions

The mean confidence ratings for the source attributions are presented in Table 3.4. Ratings ranged from 1 (guess) to 5 (certain). Due to the nature of the source attribution task, occasionally a rating was not available for a subject in a particular condition. In those few cases (less than 4%), the missing data were replaced by the mean ratings for those conditions.

The main finding of interest was the significant 3-way interaction, which demonstrated that under certain conditions of testing, confidence for incorrect source decisions may be quite strong. This interaction for test type, study format, and source attribution accuracy ($F_{(2,68)} = 6.34$, $p < .01$, $MSe = .20566$) was further investigated by two 2 x 3 ANOVAs, one for correct source attributions and one for incorrect source attributions. For the correct source attributions, confidence ratings were higher for items studied by generating (3.91) than by reading (3.63) ($F_{(1,34)} = 9.95$, $p < .01$, $MSe = .41556$) and this pattern held constant across the different test types ($F_{(2,68)} = 2.17$, $p > .05$, $MSe = .15098$).

The ratings for the incorrect source attributions exhibited a different pattern, with a significant Study X Test interaction ($F_{(2,68)} = 5.27$, $p < .01$, $MSe = .27778$). Following the procedure recommended by Kirk (1968), a

Table 3.4

Confidence Ratings for Source Attributions as a Function of Study Condition, Test Condition, and Decision Accuracy

[Experiment 2]

Correct Source Attributions	Test Condition		
	Word Identification	Generation Same	Generation Difficult
Read	3.69	3.63	3.58
Generate	3.90	4.02	3.81
Incorrect Source Attributions			
"Generate" | Read[a]	3.30	3.71	3.36
"Read" | Generate[b]	3.75	3.59	3.50

[a] words read at study that subjects reported as generated

[b] words generated at study that subjects reported as read

modified Bonferroni approach was used to evaluate the means from this interaction. The main comparisons of interest involved ratings for errors made in the word identification and generation-same conditions where the test conditions were representative of source specific information. The alpha level of .05 was divided so as to evaluate the two comparisons of interest at .02; the remainder was divided equally among the remaining comparisons. Subjects displayed the strongest confidence for their erroneous source decisions when those decisions were made in a testing format that was consistent with the error. That is, subjects who mistakenly called a read word "generated" were most confident that this decision was correct when they generated the word at test. Similarly, subjects who misattributed a generated word to a read source were most confident about this decision when it was made under conditions of word identification. Tests of means confirmed this pattern for the comparison between word identification (3.30) and generation-same (3.71) tests for words misattributed to a generate source ($t_{(1,34)}=3.18$, $p<.001$). Although the means were in the correct direction, the comparison between the word identification (3.75) and generation-same (3.59) tests for words misattributed to a read source failed to reach significance ($t_{(1,34)}=1.24$, $p=.12$). None of the other pairwise tests of means were significant when evaluated by the modified Bonferroni approach. The main effects from the analysis for incorrect source decisions for test ($F_{(2,68)}=2.69$, $p=.07$, $MSe=.32160$) and study ($F_{(1,34)}=3.44$, $p=.07$, $MSe=.39217$) were marginally significant.

Several other effects were significant in the 4-way ANOVA. Subjects were more confident for words they had generated (3.76) than for words they had read (3.54; $F_{(1,34)}=21.69$, $p<.0001$). Although the overall ratings were higher for correct source attributions (3.77) compared with incorrect source attributions (3.53; $F_{(1,34)}=18.85$, $p<.0001$), this finding was qualified by a significant interaction between source accuracy and time of test $F_{(1,34)}=3.97$, $p<.05$). Examination of this interaction revealed a trend for subjects' ratings to distinguish between correct (3.82) and incorrect (3.47) source decisions at immediate testing ($t_{(34)}=1.87$, $p<.05$, but not significant at the Bonferroni $p=.01$) but not at delay (correct=3.72, incorrect=3.60 ($t<1.00$, other ts<1.00). Finally, ratings were influenced by testing conditions ($F_{(2,68)}=4.13$, $p<.05$). Overall confidence ratings were higher in the word identification (3.66) and generation-same (3.73) conditions than in the generation-difficult (3.56) conditions. Unplanned comparisons indicated that this difference was reliable in the case of the generation-same and generation-difficult comparison ($t_{(68)}=2.80$, $p<.01$) but not for the word identification and generate-difficult comparison ($t_{(68)}=1.65$, $p>.05$). The difference between the word identification and generate-same conditions did not reach significance ($t_{(68)}=1.16$, $p>.05$).

New Item Analyses

Subjects' ability to accurately discriminate new items was assessed in a 3 (test type) X 2 (immediate or delayed test) ANOVA. Not surprisingly, subjects were better able to discriminate new items at immediate (.73) than at delayed

(.55) testing ($F_{(1,34)}$=17.14, p<.001, MSe=.04870). Neither the main effect for test type nor the Test Type X Delay interaction approached significance (Fs <1.00, ps>.05).

When subjects mistakenly judged a new item to be old, they also made a judgment about the source of that item. The percentage of "read" judgments for these new item errors was analyzed in a 3 (test type) X 2 (immediate or delayed test) ANOVA. The percentage of "read" designations across the three test types ranged from .72 to .83 at immediate testing and from .63 to .76 at delayed testing. Neither the main effects nor the interaction were significant in this analysis (Fs<1.61, ps>.05).

Discussion

The results of Experiment 2 provide additional evidence that testing conditions can influence the accuracy of source memory judgments. These influences were most clearly demonstrated for the case where subjects were less likely to correctly identify the source of words read at study when they were tested in the generation-same condition. The analysis of response bias corroborated these findings indicating that the response bias to respond "read" as the source was much reduced under the generation-same testing condition. The interpretation of the results for words generated at study is somewhat less clear. Although accuracy for generated words was reduced when those words were tested by word identification compared with the same generation task, accuracy was not significantly lower than in the generation-difficult condition with the result that subjects were not more likely to call a word "read" in the word identification condition than in the generation-difficult condition. Thus, the interpretation of the influence of word identification on source accuracy is not as straightforward as in Experiment 1. One possible explanation for the pattern of results in Experiment 2 is that subjects may be attributing generated words to read sources for different reasons in different conditions. For example, subjects may attribute words tested in the word identification condition to a read source because of the perceptual fluency afforded by the word identification test itself. On the other hand, words tested in the generation-difficult condition may be accorded a "read" status by default because subjects may be less likely to access critical source information due to the unusual task requirements of that test condition. There are two sorts of support consistent with the idea that different things may be going on with these two types of testing conditions. The first comes from the source bias data which indicated that response bias tends to be greater in the word identification than the generation-difficult condition, although the difference did not reach significance (p=.11). The second comes from an examination of confidence ratings made by subjects when they misattributed generated words to a read source. Subjects who made source errors under word identification were more confident that those decisions were correct than were subjects who made the same errors under generation-difficult conditions. This difference was significant at p=.02, although it failed to reach significance under the more stringent modified Bonferroni procedure. In sum, although it is unclear why the word identification results did not replicate the pattern of results found in Experiment 1, the inclusion of the generation-difficult condition may have played a role. The choice of an appropriate comparison condition remains an issue. The problem is pursued in Experiment 3 with a new third condition for comparison where words are presented in an auditory format.

Taken together, however, the findings from Experiments 1 (bias by word identification) and 2 (bias by generation) provide evidence that under certain conditions source judgment accuracy was reduced when words were tested in a format which required processing that was representative of the other type of source. Previously, Johnson and Raye (1981) found that errors in source memory could arise as a result of encoding that entailed operations that were representative of another type of source. The contribution of these two experiments is that they show that errors in source memory may also result from the conditions under which memory is tested. Perhaps the most striking aspect of the current findings is that source decisions were influenced by the processing that subjects performed at the time of source judgment, even though that processing was an obvious part of the testing situation. Apparently, subjects do not rely solely on information from encoding for their source decisions. Rather, source decisions appear to be influenced by the processing demands of the current situation, much as are other types of memory judgments (e.g., Jacoby, Kelley, & Dywan, 1989; Rabinowitz, 1990).

It is one thing to make an error about the source of a memory, but an important determinant of the repercussions of that error is the amount of confidence one has in that decision. Demonstrations of eyewitness errors due to misleading information (Loftus, 1979; Wells & Loftus, 1984) have focused attention on the questionable relationship between eyewitness confidence and testimony accuracy (e.g., Wells et al., 1979). This is a cause for concern given the influence that witness confidence has on jurors's perceptions of witness credibility (Lindsay et al., 1981; Wells et al., 1979). The present findings suggest that, under certain circumstances, strong confidence may be exhibited for erroneous source judgments. Although confidence ratings were not influenced by testing conditions when source judgments were correct, confidence ratings for incorrect source attributions showed a clear influence of testing condition when originally read words were misidentified as generated. Not only were subjects more likely to misidentify a word they had read as one they had generated when tested under generate testing conditions, but they also exhibited a high degree of confidence that this decision was correct. Similarly, subjects were more confident that their misattributions of generated items to a read source were correct when tested in the word identification condition than in any other condition, although this difference did not reach significance.

Previous research lead to two contradictory expectations concerning the influence of testing conditions on subjects' ability to recognize the words they had studied earlier. In the present research there was a trend for words, regardless of study format, to be more easily recognized when they were presented for testing by word identification than if they were presented in the generation-difficult format which required more effort to unscramble. In Experiment 1, subjects were less likely to recognize words they had read at study when testing required generation rather than identification or recognition. These findings are consistent with the idea that words are more likely to be judged old when they are more easily processed (Lindsay & Kelley, 1991) but does not fit well with the reports that increased effort results in higher levels of recognition (Peynircioğlu & Tekcan, 1993; Peynircioğlu & Watkins, 1988; Watkins & Peynircioğlu, 1990).

The examination of delayed testing revealed that there was a general reduction in source accuracy after a 24-hour delay. Interestingly, there was a trend for confidence ratings to distinguish between correct and incorrect source attributions at immediate testing although they did not do so at delayed testing.

In sum, the findings from Experiments 1 and 2 suggest that, in addition to the conditions of encoding, test conditions may also influence source attribution accuracy. An additional finding of interest from Experiment 2 was that under some circumstances subjects may be quite confident that their erroneous source decisions are correct.

An extension and replication of these source accuracy and confidence findings was undertaken in Experiment 3.

CHAPTER 4

Experiment 3:

Investigating the Relationship Between Memory Strength and Source Accuracy

Experiments 1 and 2 established that the conditions under which memory is tested can influence decisions about the source of information and the confidence a person may have in those source decisions. The purpose of Experiment 3 was to replicate these findings and investigate one further issue, the relationship between memory strength and source decision accuracy. In recent reviews of eyewitness research, Loftus (1991a; 1991b) observed that people are more likely to have their recollections modified when the passage of time allows the original memory to fade. Using an eyewitness testimony paradigm, Shaughnessy and Mand (1982) showed that stronger memories (created by viewing information twice rather than just once) are less susceptible to the misinformation effect. The idea that a generally weak or degraded memory trace may account for poor source memory is found in the neuropsychological literature (e.g., Meudell, Mayes, Ostergaard, & Pickering, 1985; Shimamura & Squire, 1991). Consistent with these ideas, the 24-hour delay between study and test in Experiment 2 resulted in a general decrease in accuracy for source memory.

Memory strength may be manipulated by varying the number of presentations of stimuli as Shaughnessy and Mand (1982) did in their study of eyewitness memory. Repeated presentations are inappropriate in a study of source memory, however, because repetition can influence the accuracy of source decisions and the specific influence varies as a function of the type of source decision being made. In previous research, source decision accuracy has been shown to be reduced by repeated presentations of perceived items, presumably because the cognitive operations resulting from the additional processing are judged to be representative of generated material (Johnson, Taylor, & Raye, 1977; Johnson, Raye, Wang, & Taylor, 1979). In contrast, the repetition of generated memories produces additional, elaborated cognitive operations and as a result source decision accuracy is increased (Johnson, Raye, & Durso, 1980).

An alternative approach is to assess memory strength by having subjects rate their confidence or certainty for recognition when memory is tested. People are able to monitor the contents of their memories with some degree of accuracy and judgments of confidence in the correctness of memory performance are moderately valid predictors of that performance (Hart, 1967; Tulving & Thomson, 1971).

In the present study, the relationship between memory strength and source memory accuracy was investigated in two ways. First, subjects were required to make ratings of item confidence at the time of the recognition. These ratings were correlated with both item recognition accuracy and source decision accuracy. Second, in an effort to investigate the influence of a longer delay between study and test, half the subjects were tested after a brief delay and half were tested after a 1-week delay. As in the previous studies, two study conditions were used to provide the sources for the information. Subjects studied words by reading or generating them from simple anagrams. Memory for these words was tested under three conditions: word identification, generation, and an auditory condition where the words were read to the subjects. The auditory condition was included to provide a comparatively "neutral" form of presentation that did not require processing that was indicative of an internal (cognitive operations) or external (perceptual clarity) source for memory.

In summary, subjects in Experiment 3 studied words by reading and generating them. After either a 5-minute or 1-week delay they were tested for their memory of the words under three conditions of presentation: word identification, generation, and auditory presentation. In each instance, they first judged whether they recognized the word and rated their confidence in that decision. They then decided if they had originally read or generated the word and gave a confidence rating for that decision. The role of memory strength in source attribution accuracy was assessed by correlating the ratings for item strength with source decision accuracy. Correlations were also calculated to assess the relationship between ratings of source decision confidence and source decision accuracy.

Method

Subjects and Design

Thirty-six University of Western Ontario undergraduates participated for course credit. A 2 X 2 X 3 factorial design was used. The two study conditions (read and generate) provided the different sources for the information. All subjects were tested under three testing conditions, two of which (word identification and generation) had been used in the two previous experiments. The third was an auditory testing condition where subjects were read the words and then made their judgements. The final factor, time of test (immediate or 1-week later) was manipulated between subjects.

Materials

The materials were the same as those used in Experiment 2 where they are described in detail. The full list of materials is presented in Appendix C.

Procedure

The procedure was similar to that used in Experiment 2 with the following exceptions. The procedure for the word identification task was modified slightly to eliminate the problem of unidentified items. Here subjects viewed the words for 33 ms unoer conditions of high contrast. Under these conditions virtually all of the words were identified (only 2 subjects missed one or two words, which were eliminated from further analysis). The generation-difficult condition of Experiment 2 was replaced with the auditory condition where the experimenter the read words aloud to the subject. The delay condition was extended from 24 hours to 1 week. At test, the subjects made ratings on item recognition (in addition to source decisions) using a 5-point scale (1=guess; 5=certain). These ratings provided the basis for an assessment of the relationship between rated item confidence and source attribution accuracy. At the end of testing, subjects were asked four questions to provide further information about their expectations and experience during the experiment. First, they were asked if they expected that they would have to remember the words they saw at the beginning of the session. They were also asked if they expected to be tested on the source or format of the words they studied. Then they were asked to provide a description of how they made the source decisions and to report on any strategies they used. Finally, they were asked if they felt that all the words in the word identification task were presented at the same speed. Some subjects in Experiment 2 spontaneously remarked that words in the word identification task appeared to be presented for different durations. Subjects in other experiments have reported similar experiences (e.g., Jacoby & Dallas, 1981; Witherspoon & Allar., 1985), and this question was included to determine the number of subjects who experienced this subjective change in presentation exposure. The full set of instructions is presented in Appendix E.

Results

Pooled error terms and Satterthwaite's adjustment for degrees of freedom were used in tests of simple main effects from interactions involving repeated measures (Winer, 1962). Predicted means were evaluated at an alpha level of .05 and the Bonferroni procedure was used for nonpredicted tests of means (Kirk, 1968).

Recognition Memory

The data for word recognition are presented in Table [illegible].1. As in the previous experiments, a generation effect was found for word recognition (Slamecka & Graf, 1978). Words generated at study (.83) were recognized more often than words which were simply read (.72; $F_{(1,34)}=22.28$, $p<.0001$, $MSe=.02594$).

The Test X Delay interaction ($F_{(2,\text{[illegible]})}=3.45$, $p<.05$, $MSe=.01313$) indicated that the conditions of testing influenced performance, depending on the time of test. When subjects were tested immediately after study, testing conditions had no influence on word recognition (means ranged from .77 to .80; all ts <1.00). For subjects who were tested one week later, in contrast, word recognition was lower in the generation condition (.72) than in either the word identification condition (.79) or auditory condition (.77). The word identification-generation comparison, however, was not significant using the Bonferroni procedure

Table 4.1

Word Recognition as a Function of Study Condition and Test Condition

[Experiment 3]

Study Condition	Test Condition		
	Word Identification	Generation	Auditory
Read	.73	.70	.75
Generate	.84	.82	.82

although it was significant at the $p=.05$ level ($t_{(68)}=1.99$, $p<.05$). The other comparisons were not significant ($ts \leq 1.44$, $p>.05$). The trend towards reduced recognition performance in the generation condition may be due to the relative fluency of word identification and auditory presentation compared with the generate task at one week delay. A similar, but significant, finding was observed in Experiment 1 at immediate testing. In Experiment 2, a trend was observed where words were less likely to be recognized when they were presented in a new format that required effortful processing. These findings are consistent with previous reports that words are more likely to be judged old when they are easily processed (Lindsay & Kelley, 1991).

Source Memory

The data for source memory are summarized in Table 4.2 and closely resemble those reported earlier. The most important finding is the replication of the Study X Test interaction found in Experiment 2 ($F_{(2,68)}=5.03$, $p<.05$, $MSe=.03717$). Here, as in the earlier experiment, the accuracy of the source decisions varied as a function of the conditions of testing. Subjects were less likely to report the correct source of an item if it was presented for testing in a format that was consistent with the other type of source. This held for both types of study conditions. Tests of means confirmed that subjects were more accurate in identifying the source of items they had read when tested under word identification conditions (.78) than generate conditions (.70; $t_{(117)}=2.14$, $p<.05$). The level of source attribution accuracy for the auditory testing format

Table 4.2

<u>Source Accuracy as a Function of Study Condition and Test Condition</u>

[Experiment 3]

Study Condition	Test Condition		
	Word Identification	Generation	Auditory
Read	.78	.70	.73
Generate	.48	.58	.57

(.73) fell between these two conditions and did not differ significantly from either (ts <1.25, p>.05) For words generated at study, source attribution accuracy was reduced when subjects were tested in the word identification condition (.48) but performance was better in the generate (.58) and auditory (.57) conditions, both of which exceeded the word identification condition (ts >2.24, p<.05) but did not differ from each other ($t_{(1,17)}$<1.00, p>.05). Overall, source attributions were more accurate when words had been read (.73) rather than generated (.54) at study ($F_{(1,34)}$=32.34, p<.05, MSe=.06123.), replicating the findings of Experiments 1 and 2 as well as Johnson and Raye's (1981) earlier work.

The 1-week delay in testing reduced the overall accuracy of source attributions from .70 to .57 ($F_{(1,34)}$=22.38, p<.05, MSe=.04411). A significant, but smaller (.07) reduction in source accuracy was found in Experiment 1 after a 24-hour delay.

A significant Test X Delay interaction ($F_{(2,68)}$=3.21, p<.05, MSe=.01582) revealed that the testing conditions had their influence at immediate testing. Further examination of the interaction confirmed that there were no differences among the various conditions of testing at delay (means ranged from .56 to .59; all ts <1.00). At immediate testing, however, overall source performance was poorer in the word identification (.67) and the generate conditions (.70) compared to the auditory condition (.74), although the auditory-word identification comparison reached significance at only the more lenient criterion of p=.05 and not by using the Bonferroni approach ($t_{(68)}$=1.67, $p<.05$), other $ts_{(68)}$=<1.00, $p>.05$).

In summary, the source decision results replicated the earlier findings. Source attributions were less accurate when subjects were tested in conditions that were representative of the other source. The one week delay in testing produced a greater reduction in source accuracy than the 24-hour delay in Experiment 2.

Source Bias Analyses

The data for the source bias analyses are presented in Table 4.3. Sensitivity (d') measures ranged from 1.04 to 1.13 and did not differ appreciably across the three test formats ($F<1.00$, $p>.05$). Sensitivity was reduced after the one week delay from 1.59 to .58 ($F_{(1,34)}$=15.65, p=.0004, MSe=1.74085). The interaction between test condition and time of testing was not significant ($F<1.00$, $p>.05$).

Response bias (C) for source decisions was calculated using the procedures described in Experiment 1. Negative C values indicate a bias towards answering "read" and positive C values indicate a bias towards responding "generate". Test conditions influenced subjects' bias to attribute words to a read source ($F_{(2,68)}$=3.11, $p<.05$, MSe=.43038). Specifically, subjects were more likely to call a word read when tested by word identification than generation ($t_{(68)}$=1.81, $p<.05$) or auditory presentation ($t_{(68)}$=2.39, $p<.01$). The generation-auditory difference was not significant ($t<1.00$, $p>.05$).

Table 4.3

Sensitivity and Bias Measures for Test Conditions

[Experiment 3]

	Test Condition		
Measure	Word Identification	Generation	Auditory
d'	1.04	1.13	1.09
C	-.60	-.32	-.23

Confidence Ratings for Source Decisions

The mean ratings for source decisions are displayed in Table 4.4. A significant Study X Test X Accuracy interaction ($F_{(2,68)}$=7.42, p<.05, MSe=.21727) was obtained for the ratings that subjects made for their source decisions. Two separate ANOVAs, one for ratings based on correct source decisions and one for ratings based on incorrect source decisions, were done to further investigate this interaction. Both of these analyses produced the same pattern of ratings as found in Experiment 2 for correct and incorrect source decisions.

The analyses for ratings made for correct source attributions revealed two main findings. Higher ratings were given to items that were generated (3.73) rather than read at study (3.43; $F_{(1,34)}$=13.83, p<.05, MSe=.33062), replicating the finding in Experiment 2. In addition, subjects were less confident about their source decisions after a 1-week delay (3.92 versus 3.23; $F_{(1,34)}$=10.28 p<.002, MSe=2.52999).

As in the previous experiment, the analyses for ratings made for incorrect source attributions revealed a significant Study X Test interaction ($F_{(2,68)}$=7.96, p<.001, MSe=.23151). The same procedure used in Experiment 2 was used to assess the differences among means for the confidence ratings (Kirk, 1968). The two critical comparisons between the word identification and generation test conditions for the two types of source errors were evaluated at p=.02 and the remainder was divided among the rest of the comparisons. When subjects

Table 4.4

Confidence Ratings for Source Attributions as a Function of Study Condition, Test Condition, and Decision Accuracy

[Experiment 3]

Correct Source Attributions	Test Condition		
	Word Identification	Generation	Auditory
Read	3.46	3.46	3.37
Generate	3.66	3.88	3.63
Incorrect Source Attributions			
"Generate" | Read[a]	3.09	3.44	3.27
"Read" | Generate[b]	3.35	3.09	3.02

[a] words read at study that subjects reported as generated

[b] words generated at study that subjects reported as read

made errors about the source of an item, they were more confident that these attributions were correct when the item had been presented for testing in the format that was consistent with the error. More specifically, when source errors were made for words that had been read at study, subjects were most confident that their erroneous judgments were correct when those items had been presented in a generate format (3.44) than if they had been presented in a word identification format (3.09; $t_{(130)}$=2.72, $p<.01$). Mean ratings for errors made in the auditory condition fell between the other two conditions but did not differ significantly from either of them (both ts<1.40, $p>.05$). In the case where words generated at study were incorrectly categorized as read, subjects were most confident when these words were presented in the word identification format (3.35) than either the generation (3.09; $t_{(130)}$=2.02, $p<.02$) or the auditory format (3.02) although the latter comparison was not significant under the constraints of the modified Bonferroni procedure ($t_{(130)}$=2.57, $p>.002$). Ratings made in the generate and auditory conditions did not differ ($t_{(130)}$=<1.00, $p>.05$).

A significant 3-way interaction for study, test, and time of test indicated that the pattern of ratings for incorrect source attributions differed as a function of when the subjects were tested ($F_{(2,88)}$=3.27, $p<.05$, MSe=.23151). These ratings are presented in Table 4.5. When tested immediately after study, the pattern of data is essentially the same as that for the overall ratings analysis for incorrect source decisions. The significant study by test interaction

Table 4.5

Confidence Ratings for Incorrect Source Attributions as a Function of Study Condition, Test Condition, and Time of Test

[Experiment 3]

	Test Condition		
Immediate Test	Word Identification	Generation	Auditory
"Generate" | Read[a]	3.04	3.69	3.41
"Read" | Generate[b]	3.55	3.25	3.03
1 Week Delayed Test			
"Generate" | Read[a]	3.14	3.19	3.13
"Read" | Generate[b]	3.15	2.93	3.01

[a] words read at study that subjects reported as generated

[b] words generated at study that subjects reported as read

($F_{(2,34)}$=10.87, p<.0002, MSe=.22429) was a result of subjects showing more confidence for words that were tested in the format consistent with the error. As in the combined analysis for immediate and delayed conditions, words that had been read at study but mistakenly attributed to generation, were rated more highly when tested under generate (3.69) conditions than the word identification condition (3.04; $t_{(62)}$=3.39, p <.01). Subjects were also more confident about words that had been generated at study but were misattributed to read when these words were tested under word identification (3.55) than generation conditions (3.25) although this caparison fell short of significance ($t_{(62)}$=1.57, p<.06). None of the other comparisons were significant using the Bonferroni procedure.

At delayed testing, however, there were no appreciable differences in the ratings across the conditions. The means ranged from 2.93 to 3.19 (all Fs <1.00).

Several other effects were significant from the overall 4-way ANOVA. A significant Accuracy X Delay interaction revealed that ratings distinguished between correct and incorrect source decisions at immediate but not at delayed testing ($F_{(1,34)}$=16.86, p<.0002, MSe=.34469). Confidence ratings were greater for correct (3.92) than incorrect (3.33) source decisions at immediate testing ($t_{(34)}$=3.02, p<.01), but not when testing was delayed by one week (correct=3.23, incorrect=3.09; $t_{(34)}$<1.00). This finding provides stronger evidence of the trend reported in Experiment 2. None of the other pairwise comparisons from the interaction approached significance ($ts \leq 1.30$).

A significant Study Condition X Accuracy interaction was found for the source decision confidence ratings ($F_{(1,34)}=5.69$, $p<.02$, MSe=.75024). Overall, ratings were higher for correct than incorrect source decisions but this difference was significant only for words studied under generate conditions. For words generated at study, the mean ratings for correct and incorrect source judgments were 3.73 and 3.16 respectively ($t_{(60)}=3.27$, $p<.001$). For words read at study, the ratings did not differ for correct (3.43) and incorrect (3.27) source attributions ($t_{(60)}<1.00$, $p>.05$). Additionally, ratings for source confidence were higher for words that had been generated (3.73) rather than read (3.43) at study but only when the source decisions were correct. This difference was not reliable using the Bonferroni procedure although it was significant at the $p=.05$ level ($t_{(57)}=1.77$, $p<.05$). When source decisions were incorrect, ratings for words that had been generated (3.16) did not differ appreciably from ratings for those words that had been read (3.27; $t_{(57)}<1.00$, $p>.05$).

Two main effects were significant. The delay of one week reduced the ratings from a mean of 3.63 to 3.16 ($F_{(1,34)}=4.87$, $p<.05$, $MSe=4.73215$). Finally, although qualified by the interaction, correct ratings (3.58) tended to be higher than incorrect ratings (3.21; $F_{(1,34)}=43.22$, $p<.0001$, $MSe=.34469$).

In summary, the most important finding from the analysis of confidence ratings for source decisions was the interaction between study condition, test condition and source decision accuracy. When source attributions were correct, the confidence ratings were unaffected by the conditions at testing. Ratings were higher for words that had been generated rather than read at study, and were reduced by a delay before test. When source attributions were incorrect, however, subjects were particularly confident in those decisions when testing conditions were consistent with the error. This effect was no longer apparent when testing was delayed by one week. Finally, overall ratings distinguished between correct and incorrect source decisions at immediate but not at delayed testing, reflecting the trend that was reported in Experiment 2.

Ratings of Item Confidence and Source Accuracy

Subjects provided ratings of their confidence for word recognition using a scale of 1 to 5 (1 = guessing; 5 = certain). These ratings were correlated with recognition performance and source decision accuracy using Gamma (γ) correlations (Nelson, 1984). Each of the γ correlations was tested from zero at p=.01 (Perlmutter & Myers, 1973).

Overall, subjects' confidence ratings for words correlated well with word recognition at both immediate (γ=.62) and 1-week delayed (γ=.64) testing (both ts> 8.00, p<.0001) replicating previous findings (e.g., Tulving & Thomson, 1971). The relationship between item confidence and recognition performance held for all but one of the conditions formed by the crossing of study and test conditions at both immediate and delayed testing. These correlations ranged from .34 to .85 and all differed significantly from zero (all

$\underline{t}$s> 3.18, $\underline{p}<.01$). The γ for words that had been read and tested in the immediate word failed to reach significance, although it was in the same direction as the others ($\gamma=.27$, $\underline{p}>.01$).

In contrast, no significant relationship was apparent between the confidence ratings for word recognition and overall source attribution accuracy at immediate testing ($\gamma=.11$; $\underline{t}_{(17)}=1.52$, $\underline{p}>.05$) and the two measures were negatively correlated at delayed testing ($\gamma=-.25$; $\underline{t}_{(17)}=6.23$, $\underline{p}<.0001$).

Further Gamma correlations were calculated between ratings of item recognition confidence and source attribution accuracy for each of the conditions formed by crossing the study and testing conditions at both immediate and delayed testing. These data are displayed in Table 4.6. Ratings of recognition confidence correlated with source attribution accuracy in only two of these conditions. At immediate testing, the memory confidence ratings correlated positively with source accuracy for generated ($\gamma=.48$) words tested in the auditory condition ($\underline{t}_{(13)}=3.02$, $\underline{p}<.01$). The other significant correlation occurred in the delay condition when generated words were tested in the generation condition ($\underline{t}_{(17)}=3.84$, $\underline{p}<.01$). These findings are consistent with the idea that in these conditions, accurate recognition, confidence ratings, and source judgments would be based on similar information. Generation at study provides a strong memory trace that supports later recognition and provides a basis for more confidence. Finally, source decisions for words generated at study should tend to be correct when the conditions of testing are consistent

Table 4.6

Gamma Correlations Between Item Confidence Ratings and Source Accuracy as a Function of Study Condition, Test Condition, and Time of Test [Experiment 3]

Immediate Testing	Test Condition		
	Word Identification	Generation	Auditory
Read	.40	-.04	-.22
Generate	.14	.34	.48*
1 Week Delayed Testing			
Read	.04	-.27	-.28
Generate	.04	.42*	.20

* indicates correlation significantly different from zero

(generation) or provide little competing source specific information (auditory) (Experiment 2).

The two patterns that emerged from the correlations of recognition confidence with recognition accuracy and source judgment accuracy are quite different. These data suggest that item strength, as assessed by recognition confidence ratings, predicts recognition quite well, but is not a reliable predictor of source attribution accuracy. Overall, there appears to be little relationship between rated item confidence and source memory accuracy.

Ratings of Source Confidence and Source Accuracy

Gamma correlations were calculated to assess the relationship between subjects' source decisions and their confidence ratings in those decisions. These correlations are presented in Table 4.7. At immediate testing, source ratings were reliable predictors of source accuracy for words generated at study that were tested in the generation ($\gamma=.61$, $t_{(17)}4.94$, $p<.0001$) or auditory conditions ($\gamma=.64$, $t_{(18)}=6.45$, $p<.0001$). At 1-week delayed testing, only the generate at study / generate at test condition produced a significant gamma ($\gamma=.43$ ($t_{(17)}2.97$, $p<.01$). None of the other correlations at immediate ($\gamma s=.06$ to $.37$; $ts\leq1.27$, $p>.01$) or delayed ($\gamma=-.22$ to $.17$; $ts\leq1.45$, $p>.01$) testing were significant.

The three conditions where a significant positive correlation between source decision accuracy and source confidence ratings were found shared some common characteristics. All the words had been generated at study and

Table 4.7

Gamma Correlations Between Source Confidence Ratings and Source Accuracy as a Function of Study Condition, Test Condition, and Time of Test [Experiment 3]

	Test Condition		
Immediate Testing	Word Identification	Generation	Auditory
Read	.37	.06	.16
Generate	.21	.61*	.64*
1 Week Delayed Testing			
Read	-.09	-.22	-.05
Generate	.09	.43*	.17

* indicates correlation significantly different from zero

were tested in consistent testing conditions (generation) or in a condition that was relatively "neutral" (auditory) with respect to source specifying characteristics.

In sum, ratings of source confidence do not appear to be generally reliable predictors of source accuracy. Positive correlations between source decisions and confidence in those decisions were found in only a few of the conditions.

New Item Analyses

Subjects were more accurate at identifying new items at immediate (.83) testing than after a one week delay (.58). A 3 (test type) X 2 (immediate or delayed test) ANOVA confirmed this finding ($F_{(1,34)}=18.41$, $p<.0001$, MSe=.09428). None of the other effects were significant.

Source attributions for these new item errors were examined in a 3 (test type) X 2 (immediate or delayed test) ANOVA which revealed a main effect for type of test ($F_{(2,48)}=3.40$, $p<.05$, MSe=.04739). That is, when subjects mistakenly called a new item "old", they were less likely to say they had read that word if it was presented for testing in the generation format (.69) than in either of the two other formats (.81 for word identification, .80 for auditory). These comparisons fell short of the $p=.02$ level required for the Bonferroni test, but were significant at the $p=.05$ level ($t_{(48)}\geq 1.82$, $p<05$). This pattern is consistent with the previously reported findings of Lindsay and Kelley (1991). There were no other significant effects from the ANOVA ($Fs\leq 2.39$, $p>.05$).

Answers to Post Test Questions

Most of the subjects (34 out of 36) responded that they did not expect to be tested for their memory of the words. Of the remaining two subjects, one said that they had not really thought about it, and the other replied that they thought they might have to remember a few of the words. None of the subjects expected to be asked about the format or source of the words. Subjects reported that they attempted to use a variety of strategies to remember the source of the words such as saying the word to themselves, trying to remember if it was a "difficult" word or not, imagery, and judging whether the word seemed familiar in that format. No one particular strategy was favoured.

Responses to the final question (concerning the apparent presentation duration of words for word identification) were available from 30 subjects. Responses were not available for 6 subjects at immediate test due to a printing error in the materials. These judgments were collected to assess whether subjects' beliefs about presentation duration influenced their source judgments. The effect of the word identification manipulation is demonstrated when subjects show reduced source accuracy for words generated at study. Of the 14 subjects who reported a change in duration, 8 showed the effect. Eight of the 16 subjects who reported no change also showed the effect. It has been suggested in previous research (Jacoby & Whitehouse, 1989) that subjects' memory judgments are less likely to be influenced when subjects have an alternative explanation for their current experience. Based on this suggestion, it might be expected that subjects who believed that the presentation durations varied might not be influenced in their source judgments. The results of this informal survey, however, suggest that subjects were equally likely to make source errors whether they thought the words had been presented for different durations or not.

Discussion

The results of Experiment 3 provide further evidence that subjects' source attributions may be influenced by the conditions of testing. The impact of testing conditions was clearest when words generated at study were presented by word identification for testing. Subjects were significantly less likely to correctly attribute these words to a generate source than in any other test condition. Consistent with these data, the analysis of response bias indicated that the bias to say "read" for the source of a word was strongest for the word identification condition. The pattern of source accuracy for words that had been read at study also varied across test conditions. Source accuracy was significantly reduced for these words when generation occurred at testing compared with word identification. The level of source accuracy for the auditory condition was less than in the word identification and greater than in the generation condition, as predicted, although it did not significantly differ from either. These results replicate and extend the findings of Experiments 1 and 2 and are consistent with the idea that source errors may arise when the testing conditions are representative of the alternative source.

The conditions of testing also affected the confidence that subjects expressed in their source decisions, but only for incorrect source decisions. As found in Experiment 2, confidence ratings for correct source decisions were not appreciably influenced by the conditions of testing. When source decisions were incorrect, however, subjects were most confident that their decisions were correct when the testing condition was consistent with the source error. The pattern of confidence ratings for incorrect source decisions is similar to that found in Experiment 2. In both experiments, subjects expressed the greatest confidence for their erroneous "generate" attributions when generation testing was used. In the present experiment, subjects were most confident in their incorrect attributions to a read source when word identification was done at test. This finding was significant when data for immediate and delayed subjects were considered together although it fell short of significance when immediate and delayed were considered separately. In Experiment 2, this comparison fell short of significance. Taken together, the confidence rating results of Experiment 2 and 3 indicate that under some testing conditions subjects may express strong confidence that their incorrect source decisions are correct.

Loftus (1991b) and others (e.g., Shaughnessy & Mand, 1982) have argued that stronger memories are less susceptible to distortion than weaker memories. Similar arguments have been made concerning memory for the source of information (e.g., Shimamura & Squire, 1991). The results of the present experiment indicate that, for source judgments involving internal and external sources, the situation is not so straightforward. Memory strength, as assessed by recognition confidence judgments, predicted source decision accuracy in only two of the twelve conditions investigated. Item strength, as assessed in this experiment, does not appear to be a reliable indicator of source decision accuracy.

The relationship between source decisions and confidence ratings for those decisions was also assessed. Again, it was found that subjects' confidence was not a reliable indicator of performance. A positive relationship between rated source confidence and source accuracy was found in only three conditions. A similar lack of predictive accuracy has been reported previously for the relationship between eyewitness confidence and eyewitness accuracy (e.g., Wells et al., 1979).

The effects of a 1-week delay were observed on several measures. The delay reduced source decision accuracy, although word recognition performance was not appreciably influenced by the delay. Subjects' confidence ratings distinguished between correct and incorrect source judgments at immediate testing but they did not do so at delayed testing. A similar trend was reported in Experiment 2. Finally, the confidence ratings were, in general, reduced by the delay.

In conclusion, the findings of Experiment 3 replicate and extend those from the previous two experiments. Source decision errors were made when words were presented in formats biased towards the alternative source. Under those same conditions, confidence was expressed for incorrect source decisions. Although ratings of memory strength predicted word recognition across almost all conditions, these same ratings were not reliable predictors of source decision accuracy. Ratings of source decision confidence also failed to predict source decision accuracy in most conditions. Finally, a delay in testing reduced both source accuracy and the usefulness of the source ratings in distinguishing between correct and incorrect source attributions. Source accuracy but not word recognition was appreciably reduced by this delay, providing further evidence that memory for information and its source may be influenced independently (Johnson & Raye, 1981).

CHAPTER 5

Experiment 4:

Examining the Influence of Intervening Activities on Later Source Attributions

It would seem to be relatively simple task to invoke a strategy to distinguish between an episode in which you are currently involved and one that occurred prior to it. Given this, it is a bit surprising that source errors were obtained at all in the previous experiments. Clearly the subjects were aware that the ongoing testing episode was distinct from the earlier study episode. Even so, they were not always able to distinguish between the information from study and testing and sometimes were influenced by the incorrect information when making their source decisions. Jacoby and his colleagues (e.g., Jacoby et al. 1988; 1989) as well as Rabinowitz (1989; 1990) have reported other circumstances where the conditions of testing have influenced subjects to make memory errors.

The purpose of Experiment 4 was to further investigate the influence of extraneous information on source memory judgments by adopting a methodology similar to that used in eyewitness testimony research. In a typical "misinformation" experiment, subjects viewed an accident videotape and were later presented with erroneous information about the content of that videotape. Later, the memory for the episode was tested. Compared with subjects who received no inconsistent information, subjects who saw the misleading information were more likely to claim to have seen things that were suggested in the later information.

In the present experiment, the influence of extraneous information on source decisions was investigated. The present experiment differs from the Experiments 1, 2, and 3 in that the misleading influences occurred in the period between initial study and final test rather than at the time of testing. As in the first three experiments, subjects studied words by reading or generating them. Prior to testing, the words were presented again in a format compatible with that used at study, the alternative format as at study, or not at all. Finally, either 5 minutes or 48 hours after the intervening presentation, the words were read aloud for the subjects to make recognition and source memory decisions.

In summary, this experiment examined the influence that events occurring between study and test (as opposed to influences at the time of testing) have on judgments of item recognition, source decisions, and confidence in those decisions.

Method

Subjects and Design

Thirty-six University of Western Ontario undergraduates participated for course credit or for pay. As in the previous experiments, subjects studied words under two conditions, read and generate, to provide different sources for the information. In the period intervening between study and test, those words were presented in the same format, the other format, or not at all. Finally, memory was tested by auditory presentation, either 5 minutes or 48 hours after the interfering information. Thus, the design was a 2 X 2 X 3 factorial design, with study and intervening conditions as within subject factors and time of test as a between subject factor.

Materials

The materials were identical to those used in Experiments 2 and 3. They were described in detail in Experiment 2 and presented in full in Appendix C.

Procedure

Subjects were tested individually. There were three main components to the procedure: study, intervening presentation, and testing. All subjects studied words in both conditions, received all types of intervening presentations, and were tested under the same auditory conditions. Half of the subjects were tested 5 min after the intervening presentations, and half were tested 48 hours later.

Study phase. As in the previous experiments, subjects studied words by reading or generating them. All words were pronounced aloud. The study procedure is described in Experiment 2. Once the study phase was completed, the subjects were familiarized with the word identification task using the procedure outlined in Experiment 2. The subjects were informed that these were practice items to acquaint them with the task that would be used later in the study.

Intervening presentation. During the interval between study and test, a third of each of the read and generated study words were repeated in a format that required the same processing as at study and a third of the words were repeated in a format that was inconsistent with the study condition. The final third of the words were not presented during this interval. Thus, during the intervening presentation, an item that was read at study was read again in word identification (consistent), generated under generate instructions (inconsistent) or not presented (no intervening presentation) again until the final auditory testing session where all words were presented. Words that had been generated at study were read under word identification (inconsistent), generated under generation instructions (consistent), or not presented (no intervening presentation) until the auditory test.

For the word identification task, words were shown on the monitor for 33 ms using the same procedure as in Experiment 3. Words for the intervening generation task were presented on a sheet of paper and subjects used a cardboard mask to view one word at a time. Words were presented in groupings of 7 or 8 items of each type, alternating between word identification and generation. Thirty new, unstudied words, 15 for identification and 15 for generation, were incorporated in the intervening lists. The materials were counterbalanced such that all words were presented in all three possible intervening conditions across all subjects.

Auditory test. After a delay of 5 min or 48 hours, subjects were tested for word recognition and source memory. All words from the study conditions and the new words added during the intervening task condition were included for a total of 135 words. The experimenter read these words aloud. For each word, the subject made a recognition judgment (yes/no) and provided a rating (1=guess, 5=certain) to indicate their confidence for that decision. It was emphasized that they should make their decisions based on the words they had originally read or generated and not on words they saw during the intervening tasks. For words judged to be old, subjects made a further discrimination as to whether they had read or generated that item at study. Subjects also rated their confidence in that decision (1=guess, 5=certain). All responses were recorded by the experimenter. Finally, subjects answered post test questions (numbers 1-3) that were outlined in Experiment 3.

The full set of instructions for Experiment 4 is presented in Appendix F.

Results

Pooled error terms and Satterthwaite's adjustment for degrees of freedom were used in tests of simple main effects from interactions involving repeated measures (Winer, 1962). Predicted means were evaluated at an alpha level of .05 and the Bonferroni procedure was used for nonpredicted tests of means (Kirk, 1968).

Recognition Memory

The word recognition data are presented in Table 5.1. Words generated (.71) at study were recognized more often than words read (.63) at study, replicating the generation effect found in the three previous experiments ($F_{(1,34)}=21.31$, $p<.0001$, $MSe=.01536$). In addition, the repetition of study words during the intervening period prior to testing improved later recognition ($F_{(2,68)}=4.70$, $p<.01$, $MSe=.02681$). Words identified (.68) and generated (.71) during the intervening period were recognized more often than words presented only at study and test (auditory =.63). However, only the comparison between words generated during the intervening period and those not presented during the intervening period was marginally significant using the Bonferroni criterion of $p=.02$, although it was significant by more lenient standards ($t_{(68)}=2.07$, $p<.025$; all other comparisons $ts \leq 1.30$). No other effects approached significance.

Table 5.1

Word Recognition as a Function of Study Condition and Intervening Presentation

[Experiment 4]

Study Condition	Intervening Presentation		
	Word Identification	Generation	None
Read	.65	.66	.59
Generate	.70	.76	.67

Source Memory

Table 5.2 summarizes source judgment accuracy for words that were recognized. Source accuracy varied across the different conditions formed by the study and intervening presentation combinations. The main finding of importance is the significant study by type of intervening presentation interaction, $F_{(2,34)}=15.24$, $p<.05$, MSe=.02442. The condition where no intervening presentations were made prior to testing in the auditory format can be considered a sort of baseline against which to compare the accuracy of the other conditions. When words were read at study, source decisions were less likely to be accurate when generation was required during the intervening period (.60) than if the word identification had intervened (.74) or no intervening presentation had been made (.73; both $ts \geq 3.34$, $p<.001$). When words had been generated at study, generating them again before testing increased their source accuracy relative to the no intervening presentation condition (.59 and .50 respectively, $t_{(1,34)}=2.31$, $p<.05$). When words that had been generated at study appeared in the word identification task prior to the final auditory test, source accuracy was the poorest of all three conditions (.45). Source accuracy in the word identification condition was significantly lower than in the generation condition ($t_{(1,34)}=3.59$, $p<.05$) but the word identification-no intervening comparison fell short of significance ($t_{(1,34)}=1.29$, $p=.10$).

Study conditions also influenced source accuracy. Source decisions for words read at study (.69) were more accurate than for words generated at

Table 5.2

Source Accuracy as a Function of Study Condition and Intervening Presentation

[Experiment 4]

Study Condition	Intervening Presentation		
	Word Identification	Generation	None
Read	.74	.60	.73
Generate	.45	.59	.50

study (.51), $F_{(1,34)}$=14.83, p<.0005, MSe=.11340. This effect has been found consistently in Johnson's reality monitoring work (e.g., Johnson & Raye, 1981; Johnson, Raye, Foley, & Foley, 1981) as well as in Experiments 1,2, and 3.

No reliable differences in source memory were found that could be attributed to the difference in the length of delay (either 5 min or 48 hours) between the intervening presentations and the memory test, $F_{(1,34)}$=2.69, p>.05, Mse=.02280).

Source Bias Analyses

Measures of sensitivity (d') and bias (C) are presented in Table 5.3. None of the effects from the ANOVA for d' were significant (Fs≤3.09, ps>.05).

Measures of response bias were calculated using the procedure outlined in Experiment 1. Negative C values indicate a bias towards responding "read" and positive C values indicate a bias towards responding "generate". Response biases varied depending on the type of intervening presentation ($F_{(2,68)}$=11.62, p<.0001, MSe=.33106). Subjects were less likely to attribute words to a read source if an intervening generation presentation had been made than in the other two conditions (ts≥3.61, p<.001). Although the bias to respond "read" was numerically stronger for the condition where word identification (-.56) occurred in the intervening period, this difference was not statistically greater than the read bias observed when no intervening presentation had been made (-.43; $t_{(68)}$<1.00, p>.05).

Table 5.3

Sensitivity and Bias Measures for Test Conditions

[Experiment 4]

Measure	Intervening Presentation		
	Word Identification	Generation	None
d'	.59	.71	.85
C	-.56	.06	-.43

Confidence Ratings for Source Decisions

The mean ratings for source decision confidence are displayed in Table 5.4. The ratings are based on a 5-point scale (1=guess; 5=certain). As in Experiments 2 and 3, confidence ratings for source decisions were investigated by separate ANOVAs for correct and incorrect source decisions. When correct source attributions were considered, words that were generated at study (3.77) received higher source confidence ratings at test than those that were read at study (3.47; $F_{(1,34)}=10.20$, $p<.05$, $MSe=.49308$). Similar findings were reported in Experiment 2 and 3. When the ratings for incorrect source decisions are considered, the pattern of data is similar to that reported in the previous experiments. That is, ratings were highest when the intervening conditions were consistent with the error. When subjects mistakenly attributed a read word to a generate source, they were most confident that this decision was correct when generation occurred between study and test. When source errors were made for words that had been generated at study, subjects were slightly more confident about these judgments when word identification had occurred between study and test than generation or no intervening presentation. None of these effects, however, were significant from the analysis for incorrect source attributions (all Fs<2.45, $p>.05$).

Several other effects were significant from the overall ANOVA for the ratings of source accuracy. A significant Study X Source Accuracy interaction indicated that the ratings distinguished between correct and incorrect source

Table 5.4

Confidence Ratings for Source Attributions as a Function of Study Condition, Intervening Presentation, and Decision Accuracy

[Experiment 4]

	Intervening Presentation		
Correct Source Attributions	Word Identification	Generation	None
Read	3.55	3.42	3.44
Generate	3.72	3.89	3.71
Incorrect Source Attributions			
"Generate" | Read[a]	3.31	3.61	3.41
"Read" | Generate[b]	3.43	3.37	3.24

[a] words read at study that subjects reported as generated

[b] words generated at study that subjects reported as read

attributions for words that were generated at study (correct=3.77; incorrect=3.35; $t_{(53)}=2.39$, $p<.01$) but not for words that were read at study (correct=3.47; incorrect=3.44; $t_{(53)}<1.00$, $p>.05$). In addition, for correct but not incorrect source judgments, there was a tendency to give higher source confidence ratings for words that had been generated rather than read at study. This comparison, however, was not significant using the Bonferroni adjustment although it was at the .05 level ($t_{(55)}=1.68$, $p<.05$; $t<1.00$ for the incorrect judgment comparison).

Although qualified by the interaction, the two main effects from this interaction were also significant. Ratings were higher for words generated (3.56) rather than read at study (3.46; $F_{(1,34)}=3.94$, $p<.05$, MSe=.29902). Correct source decisions received higher ratings than incorrect source decisions (3.62 and 3.40 respectively, $F_{(1,34)}=21.41$, $p<.0001$, MSe=.25964).

Ratings of Item Confidence and Source Accuracy

Gamma (γ) correlations were computed to assess the relationship between rated item confidence and accuracy for both item recognition and source attribution. Each γ was tested from zero at $p=.01$ (Perlmutter & Myers, 1973). As found in Experiment 3, item confidence ratings correlated well with recognition performance at both immediate ($\gamma=.33$) and delayed ($\gamma=.59$) testing ($ts\geq4.43$, $p<.0001$). Item ratings predicted item accuracy in all but two of the study-intervening presentation conditions at immediate and delayed testing ($\gamma s=.36$ to .59; $ts\geq2.42$, $p<.01$). A third condition, where words were generated at study and generated again prior to testing, was marginally significant (γ=.30; $t_{(17)}$=2.17, p=.02). Correlations were not significant for the read conditions with word identification (γ=.20) and no intervening (γ=.12) conditions at immediate testing (ts$\leq$1.00, p>.01).

Overall, there was no apparent relationship between rated item confidence and source attribution accuracy at immediate testing (γ=.01; $t_{(17)}$<1.00, p>.05) and the correlation was negative at delayed testing (γ=-.17; $t_{(17)}$=-3.98, p<.001). As in Experiment 3, the relationship between rated item confidence and source attribution accuracy was examined under the various conditions formed by the combination of study conditions, intervening presentation conditions, and time of test. The data are displayed in Table 5.5. Ratings of item confidence correlated with source accuracy in only a single condition. The correlation in one other condition was marginally significant. At immediate testing, a significant positive correlation between rated item confidence and source accuracy was found for words that had been generated at study and tested in the auditory condition (γ=.33; $t_{(17)}$=2.44, p<.01). At delayed testing, item confidence correlated negatively with source accuracy for the read at study with intervening word identification condition (γ=-.33), although this correlation fell short of significance at the p=.01 level ($t_{(17)}$=-2.06, p<.03). None of the other correlations differed significantly from zero (γs=-.19 to .27, ts$\leq$1.81, p>.05). In summary, ratings of item confidence provided a poor indication of source decision accuracy for the conditions in the present

Table 5.5

Gamma Correlations Between Item Confidence Ratings and Source Accuracy as a Function of Study Condition, Intervening Presentation, and Time of Test [Experiment 4]

	Intervening Presentation		
Immediate Testing	Word Identification	Generation	None
Read	.22	-.15	-.17
Generate	-.13	.25	.33*
48 Hour Delayed Testing			
Read	-.33	-.19	-.22
Generate	.23	.06	.27

* indicates correlation significantly different from zero

experiment. The one exception to this general finding was for items generated at study where no intervening presentations were made prior to testing. Although the ratings of item confidence did not generally predict source decision accuracy, they were reasonable indicators of recognition accuracy.

Ratings of Source Confidence and Source Accuracy

Gamma correlations between ratings of source confidence and source decision accuracy are presented in Table 5.6. Positive correlations were found in three of the conditions, all of which involved generating information at study. At immediate testing, the γ was .46 when an intervening generation presentation had been made ($t_{(14)}=3.01$, $p<.01$). At delayed testing, significant positive correlations were found for the word identification ($\gamma=.33$; $t_{(16)}=3.72$, $p<.001$) and the generation intervening presentations ($\gamma=.34$; $t_{(17)}=2.83$, $p<.01$). One other marginally significant correlation was observed where generated words at study were tested at immediate testing without an intervening presentation ($\gamma=.41$, $t_{(17)}=2.24$, $p=.02$). A significant negative correlation between source ratings and source accuracy was observed at immediate testing for the condition where words were read at study but a generation presentation intervened prior to test ($\gamma=-.34$; $t_{(17)}=-2.47$, $p<.01$). None of the other correlations were significant (γs$=-.08$ to .30, ts≤1.76, $p>.05$). Overall, the ratings of source confidence were not reliable predictors of source decision accuracy.

Table 5.6

Gamma Correlations Between Source Confidence Ratings and Source Accuracy as a Function of Study Condition, Intervening Presentation, and Time of Test

[Experiment 4]

	Intervening Presentation		
Immediate Testing	Word Identification	Generation	None
Read	.30	-.34*	-.01
Generate	-.01	.46*	.41
48 Hour Delayed Testing			
Read	.05	.02	-.08
Generate	.33*	.34*	.20

* indicates correlation significantly different from zero

Post Test Questions

The responses to the post test questions indicated, as in Experiment 3, that most of the subjects did not expect a memory test for the words. Only two of the 36 subjects reported that they thought their memory might be tested. None of the subjects expected to report on the source of their memories. To make their source decisions subjects reported using strategies such as repeating the word to themselves, deciding if extra work had been associated with a particular word, and checking a word for personal relevance. Other subjects indicated that the answers sometimes jumped to mind, they just knew, or that the words they had read were clearer in their mind. No one particular strategy dominated their responses.

Discussion

The main purpose of Experiment 4 was to examine the influence of extraneous information presented between study and test on later memory decisions. Loftus and her colleagues (Loftus, 1975, 1977, 1979; Loftus, Miller, & Burns, 1978; Wells & Loftus, 1984) demonstrated that misleading information presented between study and test reduces the accuracy of memory for the content of the original presentation. Although there are several differences between Loftus' misinformation procedure and the current investigation, the results of the present experiment indicate that information that occurs between original encoding and the time of testing may reduce the accuracy of source decisions. There were four main findings from Experiment 4. First, accuracy for source attributions was reduced when words were presented during the intervening period in a format that was inconsistent with their original presentation. This was demonstrated most clearly for the case where words read at study received a generation presentation prior to testing. Consistent with this observation, the bias analyses indicated a marked reduction in the tendency to respond "read" as a source attribution when an intervening generation presentation had been made. For the case where words generated at study received an intervening word identification presentation prior to test, source decisions were less accurate than in the other two conditions, although only marginally so compared to the no intervening presentation condition. The second important finding was that confidence ratings for source decisions were not influenced to the same degree as they were in Experiments 2 and 3. That is, the overconfidence that subjects displayed for their erroneous source judgments when misleading information was presented during testing was not apparent when the misleading information was presented between study and testing. A third finding, consistent with the findings in Experiment 3, was that ratings of word recognition confidence were not reliable predictors of source decision accuracy. In both Experiments 3 and 4, ratings of recognition confidence predicted recognition accuracy reasonably well, but were, in general, poor predictors of source decision accuracy. Similarly, in both studies, ratings of source confidence were not consistently reliable indicators of source decision accuracy. Finally, there was a tendency for words repeated between study and test to be better recognized than those presented only at study and test.

In summary, the results of Experiment 4 indicate that source decisions are vulnerable to distortion from interfering information that occurs between the time of study and test. These findings add to the results of Experiments 1, 2, and 3 where influences at the time of testing were found to reduce the accuracy of source decisions. The influence of misleading material on confidence ratings for erroneous source decisions appeared to be less when that information occurred between study and test (Experiment 4) than when it was part of the testing situation (Experiments 1, 2, and 3).

CHAPTER 6

General Discussion

The research in this thesis investigated factors that influence a particular type of source decision, the discrimination between memories that originate with others and memories that originate with ourselves. Three aspects of source memory were examined: source memory accuracy, confidence in source memory judgments, and the relationship between the confidence in a memory and the accuracy of its source. The experiments reported in this thesis provide evidence that under certain conditions source accuracy may be reduced by events that occur after encoding, that source confidence may be influenced by testing conditions, and that the relationship between the strength of a memory and the accuracy of a source decision is not a reliable one.

Two findings consistently found in the reality monitoring literature were observed in these studies. In all four experiments, recognition was better for words that were generated rather than read at study, indicating the typical generation effect (Slamecka & Graf, 1978; Johnson & Raye, 1981; Johnson et al., 1981). Also, as repeatedly found in previous work, source memory tended to be poorer for generated words than for read words (Johnson & Raye, 1981; Johnson et al., 1981).

The new information contributed by the present research concerns influences on source accuracy and confidence ratings for source decisions, and the relationship between recognition strength and source decision accuracy. Taken together, the results of the present experiments provide evidence that source attribution accuracy may be reduced when information is processed again in a manner that is representative of the alternative source. Source memory errors increased when the conflicting presentation was part of the actual memory test (Experiments 1, 2, & 3) as well as when it occurred between study and test (Experiment 4). Specifically, source accuracy was reduced when words generated at study were presented for word identification at testing (Experiments 1 & 3) as well as when words read at study were generated at the time of test (Experiment 2). A delay between study and test also reduced source decision accuracy (Experiments 2 & 3). In Experiment 4, a presentation intervening between study and test was found to reduce source accuracy for the condition where words read at study were repeated in a generation format prior to testing. The analyses of response bias provided further support for the source accuracy results in these four experiments.

Considered together, these results indicate that events occurring after a memory has been formed may influence decisions about how that memory was acquired. These findings are generally consistent with Johnson and Raye's

(1981) reality monitoring model and the more encompassing source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993). The reality monitoring model depicts source attribution as a decision process based on an evaluation of the memory for an episode. Much of the previous research focused on the role that information from encoding plays in the source decision process. The current research extends these previous investigations by demonstrating that two other influences may act on the source decision process: information presented between encoding and testing and information presented at the time of the test.

When Johnson et al. (1993) recently incorporated reality monitoring into their larger source monitoring framework, they also provided an elaborated description of the types of judgment processes involved in source attribution. According to their framework, most source decisions are made rapidly without much awareness of the source decision process. At other times, source decisions may involve a more strategic and deliberate retrieval of source information and the use of additional information such as prior knowledge to assess the plausibility of the decision. Johnson et al. refer to these as heuristic and systematic based source decisions respectively. The source decisions made in the present experiments were likely to have been heuristic type decisions; subjects were required to make source decisions relatively quickly without much time for reflection. Under these conditions, it was observed that source decisions were often subject to bias when inconsistent information was presented between study and test or at the time of test. There is evidence from Lindsay & Johnson (1989b) that memory errors may be reduced when subjects are induced to adopt a more systematic rather than heuristic approach to memory judgments. They found a reduction in the "misinformation effect" in the eyewitness testimony paradigm when subjects were directed to think about source of information at the time of test. Although it was not a focus of the present research, this finding raises the possibility that if subjects could invoke a more strategic approach to their source decisions, either voluntarily or at the experimenter's suggestion, source attribution errors may be reduced.

Two claims made in the reality monitoring model received additional support from the current research. Johnson and Raye (1981; also Johnson et al., 1993) have argued that, because the two types of memory, memory for occurrence and memory for origin, may be influenced differently by the same variables, source memory requires a model separate from that of other memory. Consistent with this claim, it was found in all four of the present experiments that generating words at study resulted in improved recognition but poorer source memc y. The second point concerns the source attribution process itself. Johnson and Raye depict the source decision process, at least for this particular type of source decision, as not simply a retrieval of a source label but rather as a decision based on an examination of the characteristics of information in memory. The findings from the present research are not easily accommodated by a memory framework that characterizes the source decision process as reading information from a label or tag associated with a memory (e.g., Anderson, 1983). In particular, the findings from the first three experiments present a problem for this type of account. While a "tag confusion" account is plausible for the errors arising from intervening presentations, it is unlikely that source confusions arising from the current testing conditions could be accounted for in this way. The distinction the subject must make is between an experience in the past and one in which they are currently involved while the stimulus is still present.

The attributional analysis of memory framework developed by Jacoby and his colleagues (e.g., Jacoby, Kelley, & Dywan, 1989; Kelley & Jacoby, 1990) emphasizes the role of subjective experience in remembering and research has demonstrated these influences for many types of memory judgments. For example, if a word is easily read relative to others during a recognition task, a subject is likely to say that they had read it earlier. According to Jacoby et al., the ease with which the word is read relative to others results in a feeling of familiarity and which in turn leads to the recognition judgment. The subject makes an attribution about memory based on the subjective experience of reading the word. An attribution process underlies the source monitoring process as well, but Johnson has focused on attributions made on the basis of information available from the encoding episode. The pattern of source decision errors in the present research is consistent with the idea that subjects are sometimes influenced by their current experience when making source attributions, much in the way that subjects in the Jacoby studies were for their various memory judgments.

In eyewitness testimony research (e.g., Loftus, 1975), subjects claimed to have witnessed things they actually only heard or read about. The results of the present studies indicate that, under certain conditions, subjects may also claim to have generated information when they have simply read it, or alternatively, to have read information they have generated. The findings of the present studies indicate that memory for the source of information, like memory for the target information itself, may be altered by later events.

A second focus of interest concerned the confidence displayed by subjects for their source attributions. In eyewitness testimony research, interest has centred on the relationship between the accuracy of eyewitness reports and the confidence that subjects display in their memories (e.g., Wells, Lindsay, & Ferguson, 1979). This relationship is of interest in itself and also because of the influence that perceived witness confidence has on juries (Lindsay, Wells, & Rumpel, 1981; Wells, Lindsay, & Ferguson, 1979). The amount of confidence that a person has in a particular memory has practical consequences as well: Confidence is a good predictor of whether or not a person will act on a memory (Jacoby et al., 1989). People, however, are often overconfident in evaluating the correctness of their knowledge (Koriat, Lichtenstein, & Fischhoff, 1980).

The results from the present research indicate that when source decisions were correct, confidence ratings were higher for words that were generated compared with words that were read at study (Experiments 2 & 3), although overall source accuracy was greater for words that had been read than those that had been generated. When confidence ratings for incorrect source decisions were considered, an interesting pattern emerged. Subjects were most confident that their incorrect source attributions were correct when the conditions at test were consistent with the source error. This pattern was observed for both types of errors in Experiment 3 and when words were misattributed to a generate source in Experiment 2.

In Experiment 4, where the effect of intervening influences that occur between study and test was examined, confidence ratings for correct source decisions acted much as those in the previous experiments in that subjects were more confident that source decisions were correct for words they had generated at study than for words they had read at study. When ratings for incorrect source decisions were considered, however, intervening presentations did not influence the source confidence ratings to the same degree as was observed in the previous two experiments. In Experiments 2 and 3, subjects were especially confident in their incorrect source decisions when those items were presented for testing in a format that was consistent with the error. That this overconfidence was apparent only when the misleading information was present at test and not when it was presented between encoding and testing, suggests that the immediate presence of misleading information may influence confidence. In an examination of the processes underlying confidence in many types of judgments, Koehler (1991) has reported that when a task requires a person to treat a hypothesis as though it were true, confidence increases in that hypothesis. Consistent with this account, under certain conditions subjects expressed the greatest confidence that their incorrect source attributions were correct when the test conditions were representative of the source to which they attributed the word.

Other confidence findings indicated that the source ratings distinguished between correct and incorrect source decisions at immediate but not at one week delayed testing (Experiment 3). A similar, but nonsignificant, trend was observed in Experiment 2, which used a 24-hour delay. A general reduction was observed for the ratings after a one week delay (Experiment 3).

Correlations were calculated to assess the relationship between rated source confidence and source decision accuracy for Experiment 3 where testing conditions were manipulated and Experiment 4 where intervening conditions were manipulated. For both experiments, significant correlations were found in only a few of the conditions, indicating that ratings of source confidence were not, in general, good predictors of source decision accuracy.

In sum, confidence for source decisions appears to be subject to a number of influences - the conditions of study, the conditions at test, and whether a significant delay has passed between study and test. Others have reported that confidence ratings are subject to other influences such as the speed of retrieval from memory (Kelley & Lindsay, 1993; Nelson & Narens 1990). An additional contribution from the present work is that, under particular testing conditions, subjects may express greater confidence that their incorrect source attributions are correct.

A third issue of interest concerned the relationship between memory for a word and the accuracy of memory for its source. Evidence from a variety of studies supports the idea that as the memorability of a particular episode increases, so increases the memorability of a given aspect of that episode. Hayman & Rickards (1993) reported that recall of study modality (visual or auditory) was facilitated by factors such as elaboration of meaning that also increased recall of the material itself. Others have suggested that there is a relationship between memory strength and the accuracy of source judgments for those memories (Meudell et al., 1985; Shimamura & Squire, 1991). In the present research the idea that strong memories are accompanied by correct source attributions was assessed by correlating rated item confidence and source memory accuracy. Although ratings of memory strength (as indicated by confidence in recognition) generally predicted word recognition performance, strong memories were not necessarily associated with accurate source attributions (Experiments 3 & 4). An additional finding supporting the independence of word recognition and source decision was the observation that while delays of 24 hours or 1 week reduced source accuracy, there was no appreciable reduction in recognition memory over these periods. It appears that under some circumstances there may be a tendency for memories and specific characteristics of those memories to be affected similarly, but the relationship does not necessarily hold for the type of memory distinctions under investigation in the present research. These findings are consistent with the Johnson and Raye (1981; also Lindsay & Johnson, 1991) observations that memory for occurrence and origin do not necessarily respond to the same variables.

Finally, the differences between memories that occur in everyday life and memories created for investigation in a laboratory setting must be addressed. Memories for real world experiences are often much richer (e.g., may contain affective components) and more detailed than memories for words studied in the lab. The task of recalling a list of words and making judgments about them is also somewhat removed from our everyday experience of remembering. However, given these limitations, laboratory investigations of cryptomnesia, eyewitness testimony, and source memory may provide useful insights into the real world memory phenomena and provide the bases for further research.

There are many other issues of interest that the research in this thesis does not address such as the intentions of the subject, different types of source memory, and different types of internal and external sources for memories. For example, subjects may be able to invoke specific strategies to aid in their source decisions. This issue was raised in Experiment 1, where it was found that subjects' source decisions for words they read at study were not influenced by testing conditions when those subjects experienced only one testing condition. This possibility remains speculative, however, since that research was not designed to address that issue and alternative interpretations are not ruled out.

A Direction for Future Research

In 1985, Tulving introduced the terms "remember" and "know" to describe differences in the subjective experience of remembering. The hallmarks of a memory that is "remembered" are a subjective feeling of having seen the item during the study episode and a conscious recollection that it was on the study list. These "remembered" memories may be quite detailed and vivid. In contrast, a memory that subjects judge as "know" refers to an item they can tell was on the study list (often with certainty) but cannot recollect the actual occurrence. Recent work indicates that these judgments are not made solely on the basis of confidence (Rajaram, 1993). Gardiner and his colleagues (1988; Gardiner & Java, 1990, 1991; Gardiner & Parkin, 1990; also Rajaram, 1993) have investigated these two types of memory judgments in tests of recognition memory and both types of judgments contribute to performance on what are considered to be explicit tests of recollection. A similar approach could be usefully applied to the investigation of factors that influence source memory decisions. Source decisions also appear to be based on more than one type of information.

Conclusions

Linton (1982) has observed that difficulty in remembering the origin of memories is among the most commonly encountered memory problems. Judgments of origin are known to underlie many important cognitive activities (Johnson, Hashtroudi, & Lindsay, 1993). An understanding of the processes by which the sources of our memories are identified and specification of the factors that influence the accuracy and confidence of such decisions has practical as well as theoretical significance. The studies reported in this thesis represent a modest attempt to address these issues.

We commonly discuss events with others or ruminate on past events. The present research suggests that, in some instances, these activities may compromise the accuracy of our judgments about the origins of memories and influence our confidence in those judgments.

APPENDIX A

Materials Used in Experiment 1

LIST A

author
bedroom
welfare
battery
signal
anxiety
theorem
triumph
unusual
ritual
slavery
inquiry
mixture
journey
version
dispute
passion
transit
payment
finance
gallery
tractor
caution
heaven
utility
witness
uniform
worship
highway
lecture

LIST B

veteran
loyalty
impulse
flower
engine
package
edition
ballet
kingdom
revenue
minimum
library
liberty
disease
dancer
carbon
poverty
tactics
affair
leather
arrival
harmony
venture
purchase
gesture
neglect
garden
routine
passage
fashion

LIST C

barrel
breath
warfare
insight
salary
dilemma
wisdom
muscle
belief
embassy
textile
oxygen

female
marble
unknown
cabinet
message
tourist
remark
horizon
emotion
warning
avenue
realism
mustard
radical
autumn
tangent
silence
canvas
shadow
feature

USED IN WORD IDENTIFICATION PRACTICE

theatre
warrant
billion
benefit
aspect
despair
parade
jungle
weather
stadium
protein
miracle
luxury
finger
network
surplus
branch
article
illness
soldier
camera
vehicle

APPENDIX B

Instructions Used in Experiment 1

INSTRUCTIONS FOR WORD IDENTIFICATION PRACTICE

In this part of the study you will read words from the computer monitor. The words will be shown very briefly and your task is to simply read them aloud. For each word that is presented the following sequence will take place. First you will see a "+" on the screen to help you know where to look. This will be followed by a series of nonsense letters "&&&&&&&&". Then a word will flash on the screen. Read this word aloud when you see it. The word will be followed by another series of nonsense letters. This sequence will be repeated for each of the words that is presented. I will initiate the presentation of each word by hitting the space bar on the keyboard. The words come up quickly on the monitor, so just do your best to read them aloud. These words will be for practice to familiarize you with the procedure which we will use later in the study. Any questions?

STUDY PHASE

A series of words will be presented on computer monitor. Your task will be to read each word aloud as it is shown. Two types of words will appear. Some of these words will be presented in a normal fashion such as "HOUSE". Other words will be presented with the first two letters reversed such as "OHUSE". Your task with these words is to unscramble the letters and pronounce the word as it should be. That is, you are to read it as a proper english word. In sum, two types of words will be shown to you on the screen: words printed normally and words with first two letters reversed. Your task in each instance is to say the correct english word aloud. The computer controls the presentations of these words and they will continue to come up in a series. A message will appear when the list is completed. Do you have any questions?

INSTRUCTIONS FOR STATE GENERATION TASK

This time your task is to write down the names of all the states that make up the United States of America. You may write them in any order and use any strategy you wish to try and remember them. You have 5 minutes for this task.

TESTING PHASE

WORD IDENTIFICATION INSTRUCTIONS: For this part of the study, you will be presented with words and make some judgments about them. The words will be presented very quickly on the monitor as they were in the task you did when you first arrived. All the words you see will be presented as normal words (e.g., HOUSE). For each word you see your task is to read it aloud. After you have said the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you read aloud earlier, the one with words you read and words you unscrambled. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list.

To recap, words will be presented on the monitor and your task is to read each word aloud and decide if it was one you saw earlier when you read and unscrambled words on the monitor. For words that you recognize, you'll decide whether you read that word or had to unscramble it earlier. I will write down all your responses.

GENERATE TEST INSTRUCTIONS: For this part of the study, you will be presented with words and make some judgments about them. The words will be presented on a printed page and you will use this mask that allows you to view only one word at a time. All the words you see will be presented with their first two letters reversed (e.g., OHUSE).

For each word you see, your task is to read it aloud (as the word it should be in english). After you have said the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you read aloud earlier, the one with words you read and words you unscrambled. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list.

To recap, words will be presented on the paper for you to read. Your task is to unscramble each word, say it aloud, and decide if it was one you saw earlier when you read and unscrambled words on the monitor. For words that you recognize, decide whether you read that word or had to unscramble it earlier. I will write down all your responses.

RECOGNITION TEST INSTRUCTIONS: For this part of the study, you will be presented with words and make some judgments about them. The words will be presented on a printed page using this mask that allows you to view only one word at a time. All the words you see will be presented as normal words (e.g., HOUSE).

For each word you see, your task is to read it aloud. After you have said the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you read aloud earlier, the one with words you read and words you unscrambled. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list.

To recap, words will be presented on the paper for you to read. Your task is to repeat each word aloud and decide if it was one you saw earlier when you read and unscrambled words on the monitor. For words that you recognize, decide whether you read that word or had to unscramble it earlier. I will write down all your responses.

APPENDIX C

Materials Used in Experiments 2, 3, & 4

LIST A

radical
tractor
gallery
avenue
author
bedroom
welfare
remark
battery
anxiety
breath
cabinet
theorem
signal
canvas
tangent
slavery
ritual
mixture
mustard
lecture
highway
utility
warning
emotion
salary
worship

uniform
witness
autumn
heaven
caution
finance
payment
inquiry
belief
dispute
version
unknown
triumph
unusual
journey
transit
passion
muscle

LIST B

marble
wisdom
veteran
flower
barrel
loyalty
engine
ballet
library
kingdom
venture
tourist
impulse
fashion
passage
routine
garden
package
edition
affair
tactics
poverty
female
comfort
insight
revenue
neglect

gesture
realism
horizon
oxygen
message
harmony
correct
arrival
leather
minimum
carbon
dilemma
disease
dancer
textile
embassy
silence
liberty

LIST C

ancient
typical
plenty
achieve
estate
formal
liquid
atomic
curious
pocket
contain
lawyer
storage
visible
courage
divorce
colony
butter
climate
legend
fortune
saddle
tribute
crystal
plaster
prairie
guitar
servant
tobacco
candle
imagine
formula
tension
comedy
chicken
dignity
logical
healthy

massive
ceiling
powder
holiday
buffalo
cottage
exhibit

APPENDIX D

Instructions Used in Experiment 2

INSTRUCTIONS FOR WORD IDENTIFICATION PRACTICE

In this part of the study you will read words from the computer monitor. The words will be shown very briefly and your task is to simply read them aloud. For each word that is presented the following sequence will take place. First you will see a "+" on the screen to help you know where to look. This will be followed by a series of nonsense letters "&&&&&&&&". Then a word will flash on the screen. Read this word aloud when you see it. The word will be followed by another series of nonsense letters. This sequence will be repeated for each of the words that is presented. I will initiate the presentation of each word by hitting the space bar on the keyboard. The words come up quickly on the monitor, so just do your best to read them aloud. These words will be for practice to familiarize you with the procedure which we will use later in the study. Any questions?

Occasionally I will adjust the monitor controls during the procedure.

[6 words presented @ 66ms, 9 words @ 33 ms, last 9 set manually]

STUDY PHASE

A series of words will be presented on computer monitor. Your task will be to read each word aloud as it is shown. Two types of words will appear. Some of these words will be presented in a normal fashion such as "HOUSE". Other words will be presented with the first two letters reversed such as "OHUSE". Your task with these words is to unscramble the letters and pronounce the word as it should be. That is, you are to read it as a proper english word. In sum, two types of words will be shown to you on the screen: words printed normally and words with first two letters reversed. Your task in each instance is to say the correct english word aloud. The computer controls the presentations of these words and they will continue to come up in a series. A message will appear when the list is completed. Do you have any questions?

[Immediate or delayed testing]

TESTING PHASE

For this part of the study, you will read words and make some judgments about them. Some of the words will be presented very quickly on the monitor. You will read some other words from a printed page using this mask that allows you to view only one word at a time. There are two types of words on the printed page. Some of them are printed with the first two letters reversed [HCANGES]. Others are scrambled in a more complex fashion with the letters rearranged according to the following rule [GNESAHC = CHANGES]. [show card with complex scrambled rule, 5467321, and explain how it works] This card will remain in view for you to consult.

For each word you see, regardless of how it is presented, your task is to read it aloud as the word it should be in english.

After you have read the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you read aloud earlier, the one with words you read and words you unscrambled. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list. Finally, I'd like you to give me an idea about how certain you are about the decision about how you saw the word earlier. A rating of 5 would mean you are certain about your answer; a rating of 1 would mean you are guessing. Feel free to use the full range of numbers to describe how confident you are about your response.

To recap. words will be presented on the monitor and on the paper for you to read. They will be presented in 3 different formats: very quickly, with the first 2 letters reversed, or with the letters scrambled according to the rule presented on the card. Regardless of the presentation format, your task is to read each word aloud and decide if it was one you read earlier when you read and unscrambled words on the monitor. For words that you recognize, decide whether you read that word or had to unscramble it earlier. Then give me your rating for how confident you feel about that decision. I will write down all your responses.

APPENDIX E

Instructions Used in Experiment 3

INSTRUCTIONS FOR WORD IDENTIFICATION PRACTICE

In this part of the study you will read words from the computer monitor. The words will be shown very briefly and your task is to simply read them aloud. For each word that is presented the following sequence will take place. First you will see a "+" on the screen to help you know where to look. This will be followed by a series of nonsense letters "&&&&&&&&". Then a word will flash on the screen. Read this word aloud when you see it. The word will be followed by another series of nonsense letters. This sequence will be repeated for each of the words that is presented. I will initiate the presentation of each word by hitting the space bar on the keyboard. The words come up quickly on the monitor, so just do your best to read them aloud. These words will be for practice to familiarize you with the procedure which we will use later in the study. Any questions?

STUDY PHASE

A series of words will be presented on computer monitor. Your task will be to read each word aloud as it is shown. Two types of words will appear. Some of these words will be presented in a normal fashion such as "HOUSE". Other words will be presented with the first two letters reversed such as "OHUSE". Your task with these words is to unscramble the letters and pronounce the word as it should be. That is, you are to read it as a proper english word. In sum, two types of words will be shown to you on the screen: words printed normally and words with first two letters reversed. Your task in each instance is to say the correct english word aloud. The computer controls the presentations of these words and they will continue to come up in a series. A message will appear when the list is completed. Do you have any questions?

[Immediate or delayed testing]

TESTING PHASE

For this part of the study, you will be presented with words and make some judgments about them. Some of the words will be presented very quickly on the monitor. You will read some other words from a printed page using this mask that allows you to view only one word at a time. Finally, I will read some words. So there are three sorts of words: words that will be presented on the monitor, words you will read from the sheet on your desk, and words you will hear me read.

For each word you see, regardless of how it is presented, your task is to repeat it aloud (as the word it should be in english). After you have said the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you read aloud earlier, the one with words you read and words you unscrambled. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list. Finally, I'd like you to give me an idea about how certain you are about the decision about how you saw the word earlier. A rating of 5 would mean you are certain about your answer; a rating of 1 would mean you are guessing. Feel free to use the full range of numbers to describe how confident you are about your response.

To recap, words will be presented on the monitor, on the paper for you to read, or I will read them for you to hear. Regardless of the presentation format, your task is to repeat each word aloud and decide if it was one you saw earlier when you read and unscrambled words on the monitor. For words that you recognize, decide whether you read that word or had to unscramble it earlier. Then give me your rating for how confident you feel about that decision. I will write down all your responses.

POST TEST QUESTIONS

1. Did you expect to be asked to remember the words you saw at the beginning of the session?
2. Did you expect to have to report the source (format) of the words you studied?
3. Please describe any strategies you used to try and remember the source of the words.
4. Think about when you did the Word Identification task during this last part of the experiment. Did some of these words seem to presented more quickly than others or did they all appear at about the same speed?

APPENDIX F

Instructions Used in Experiment 4

INSTRUCTIONS FOR WORD IDENTIFICATION PRACTICE

In this part of the study you will read words from the computer monitor. The words will be shown very briefly and your task is to simply read them aloud. For each word that is presented the following sequence will take place. First you will see a "+" on the screen to help you know where to look. This will be followed by a series of nonsense letters "&&&&&&&&". Then a word will flash on the screen. Read this word aloud when you see it. The word will be followed by another series of nonsense letters. This sequence will be repeated for each of the words that is presented. I will initiate the presentation of each word by hitting the space bar on the keyboard. The words come up quickly on the monitor, so just do your best to read them aloud. These words will be for practice to familiarize you with the procedure which we will use later in the study. Any questions?

STUDY PHASE

A series of words will be presented on computer monitor. Your task will be to read each word aloud as it is shown. Two types of words will appear. Some of these words will be presented in a normal fashion such as "HOUSE". Other words will be presented with the first two letters reversed such as "OHUSE". Your task with these words is to unscramble the letters and pronounce the word as it should be. That is, you are to read it as a proper english word. In sum, two types of words will be shown to you on the screen: words printed normally and words with first two letters reversed. Your task in each instance is to say the correct english word aloud. The computer controls the presentations of these words and they will continue to come up in a series. A message will appear when the list is completed. Do you have any questions?

INTERVENING PRESENTATIONS PHASE

Now I would like you to read some other words aloud. These words will be presented in 2 different formats. Some of the words will be presented very quickly on the monitor. Others you will read from the sheet on the desk. These will have the first two letters reversed (OHUSE) and your task is to read them aloud as the words they should be in english.

[immediate or 48 hour delayed testing]

TESTING PHASE

For this part of the study, you will hear words and make some judgments about them. After you have repeated the word aloud, I will ask you a series of questions about it. First I will ask you to make a judgment (yes/no) as to whether or not you recognize the word from the list you saw aloud earlier, the one with words you read and words you unscrambled. Please be ensure that you are only saying "YES" to words from that original list, not from any of the tasks you have done since then. Next for words you do recognize, I'd like you to make a further decision about whether you think you read that word or unscrambled it in that earlier list. Finally, I'd like you to give me an idea about how certain you are about the decision about how you saw the word earlier. A rating of 5 would mean you are certain about your answer; a rating of 1 would mean you are guessing. Feel free to use the full range of numbers to describe how confident you are about your response.

To recap, words will be presented on the monitor and on the paper for you to read. For words that you recognize, decide whether you read that word or had to unscramble it earlier. Then give me your rating for how confident you feel about that decision. I will write down all your responses.

POST TEST QUESTIONS

1. Did you expect to be asked to remember the words you saw at the beginning of the session?
2. Did you expect to have to report the source (format) of the words you studied?
3. Please describe any strategies you used to try and remember the source of the words.

REFERENCES

Anderson, J.R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261-295.

Anderson, R.E. (1984). Did I do it or did I only imagine doing it? *Journal of Experimental Psychology: General*, *113*, 594-613.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*, 610-632.

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

Bowers, K.S., & Hilgard, E. (1988). Some complexities in understanding memory. In H.M. Pettinati (Ed.), *Hypnosis and memory* (pp. 3-18). New York: Guilford Press.

Brown, A.S., & Halliday, H.E. (1991). Cryptomnesia and source memory difficulties. *American Journal of Psychology*, *104*, 475-490.

Brown, A.S., & Murphy, D.R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 432-442.

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

Cohen, G, & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and eyewitness testimony. *Psychology and Aging*, 4, 10-17.

Cole, W.G., & Loftus, E.F. (1979). Incorporating new information into memory. American Journal of Psychology, 92(3), 413-425.

Craik, F.I.M., Morris, L.W., Morris, R.G., & Loewen, E.R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. Psychology & Aging, 5, 148-151.

Durso, F.T., & Johnson, M.K. (1980). The effects of orienting tasks on recognition, recall, and modality confusions of pictures and words. Journal of Verbal Learning and Verbal Behavior, 19, 416-429.

Dywan, J., & Jacoby, L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. Psychology and Aging, 5, 379-387.

Finke, R.A., Johnson, M.K., & Shyi, G.C.W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. Memory and Cognition, 16, 133-137.

Foley, M.A., & Johnson, M.K. (1985). Confusion between memories for performed and imagined actions. Child Development, 56, 1145-1155.

Foley, M.A., Johnson, M.K., & Raye, C.L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. Child Development, 54, 51-60.

Freud, S. (1972). The psychopathology of everyday life. (A. Tyson, Trans.). J. Strachey (Ed.). London: Penguin Books. (Original work published 1901).

Gardiner, J.M. (1988). Functional aspects of recollective experience. Memory & Cognition, 16, 309-313.

Gardiner, J.M., & Java, R.I. (1990). Recollective experience in word and nonword recognition. Memory and Cognition, 18, 23-30.

Gardiner, J.M., & Java, R.I. (1991). Forgetting in recognition memory with and without recollective experience. Memory and Cognition, 19, 617-623.

Gardiner, J.M., & Parkin, A.J. (1990). Attention and recollective experience in recognition memory. Memory and Cognition, 18, 579-583.

George Harrison guilty of plagiarizing, subconsciously, a '62 tune for a '70 hit.

(1976, September 8). New York Times. p. 50.

Goldstein, P. (1989). Copyright: Principles, law, and practice (Vol. II, pp. 161-164). Boston: Little, Brown and Company.

Hart, J.T. (1967). Memory and the memory monitoring process. Journal of Verbal Learning and Verbal Behavior, 6, 685-691.

Hashtroudi, S., Johnson, M.K., & Chrosniak, L.D. (1989). Aging and source monitoring. Psychology and Aging, 4, 106-112.

Hayman, C.A.G., & Rickards, C. (1993). A dissociation in the effects of study modality on tests of implicit and explicit memory. Manuscript submitted for publication.

Jacoby, L.L., Allan L.G., Collins, J.C., & Larwill, L.K. (1988). Memory influences subjective experience: Noise judgments. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 240-247.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General, 110, 306-340.

Jacoby, L.L., Kelley, C.M., & Dywan, J. (1989). Memory attributions. In H.L. Roediger and F. I. M. Craik (Eds.), Varieties of memory and consciousness: Essays in honor of Endel Tulving (pp. 391-422). Hillsdale, NJ: Erlbaum.

Jacoby, L.L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. Journal of Experimental Psychology: General, 118, 126-135.

Jacoby, L.L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. Journal of Experimental Psychology: General, 118, 115-125.

Janowsky, J.S., Shimamura, A.P., Kritchevsky, M., & Squire, L.R. (1989). Source memory impairment in patients with frontal lobe lesions. Neuropsychologia, 7, 1043-1056.

Johnson, M.K. (1985). The origin of memories. In P.C. Kendall (Ed.), Advances in cognitive-behavioral research and therapy, (Vol. 4, pp. 1-26). San Diego, CA: Academic Press.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. Psychological Bulletin, 114, 3-28.

Johnson, M.K., & Raye, C.L. (1981). Reality monitoring. Psychological Review, 88, 67-85.

Johnson, M.K., Raye, C.L., & Durso, F.T. (1980). Reality monitoring: Second perceptions and thoughts. Bulletin of the Psychonomic Society, 15, 402-404.

Johnson, M.K., Raye, C.L., Foley, H.J, & Foley, M.A. (1981). Cognitive operations and decision bias in reality monitoring. American Journal of Psychology, 94, 37-64.

Johnson, M.K., Raye, C.L., Wang, A.Y., & Taylor, T.H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. Journal of Experimental Psychology: Human Learning and Memory, 5, 229-240.

Johnson, M.K., Taylor, T.H., & Raye, C.L. (1977). Fact and fantasy: The effects of internally generated event on the apparent frequency of externally generated events. Memory and Cognition, 5, 116-122.

Kelley, C.M., & Jacoby, L.L. (1990). The construction of subjective experience: Memory attributions. Mind & Language, 5, 49-68.

Kelley, C.M., Jacoby, L.L., & Hollingshead, A. (1989). Direct versus indirect tests of memory for source: Judgments of modality. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 1101-1108.

Kelley, C.M., & Lindsay, D.S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. Journal of Memory and Language, 32, 1-24.

Kirk, R.E. (1968). Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole.

Koehler, D.K. (1991). Explanation, imagination, and confidence in judgment. Psychological Bulletin, 110, 499-519.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Kučera, H., & Francis, W.N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Lindsay, D.S. (1987). Whence comes this memory? Unpublished doctoral dissertation, Princeton Unive‵‵˙., Princeton, NJ.

Lindsay, D.S., & Johnson, M.K. (1989a). The reversed eyewitness suggestibility effect. Bulletin of the Psychonomic Society, 27, 111-113.

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

Lindsay, D.S., & Johnson, M.K. (1991). Recognition memory and source monitoring. Bulletin of the Psychonomic Society, 29, 203-205.

Lindsay, D.S., & Kelley, C.M. (1991, June). Ease of generation during recall induces a feeling of remembering. Poster presented at the meeting of the American Psychological Society, Washington, D.C.

Lindsay, R.C.L., Wells, G.L., & Rumpel, C.M. (1981). Can people detect eyewitness-identification accuracy within and across situations? Journal of Applied Psychology, 66, 79-89.

Linton, M. (1982). Transformations of memory in everyday life. In U. Neisser (Ed.), Memory observed: Remembering in natural contexts (pp. 77-91). San Francisco: Freeman.

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

Loftus, E.F. (1977). Shifting human color memory. Memory & Cognition, 5, 696-699.

Loftus, E.F. (1979). Eyewitness testimony. Cambridge, Massachusetts: Harvard University Press.

Loftus, E.F. (1991a). Made in memory: Distortions in recollection after misleading information. In The psychology of learning and motivation (Vol. 27, pp. 187-215). New York: Academic Press.

Loftus, E.F. (1991b). Memory distortions. Canadian Psychology, 32, 645-647.

Loftus, E.F., & Hoffman, H.G. (1989). Misinformation and memory: The creation of new memories. Journal of Experimental Psychology: General, 118, 100-104.

Loftus, E.F., Korf, N.L., & Schooler, J.W. (1989). Misguided memories: Sincere distortions of reality. In J. Yuille (Ed.), Credibility assessment: A theoretical and research perspective (pp.155-174). Kluwer: Dordrecht.

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

Marsh, R.L., & Bower, G.H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 673-678.

Macmillan, N.A., & Creelman, C.D. (1991). Detection theory: A user's guide. New York: Cambridge University Press.

McIntyre, J.S., & Craik, F.I.M. (1987). Age differences in memory for item and source information. Canadian Journal of Psychology, 41, 175-192.

Meudell, P.R., Mayes, A.R., Ostergaard, A., & Pickering, A. (1985). Recency and frequency judgements in alcoholic amnesics and normal people with poor memory. Cortex, 21, 487-511.

Morton, J., Hammersley, R.H., & Bekerian, D.A. (1985). Headed records: A model for memory and its failures. Cognition, 20, 1-23.

Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 95, 109-133.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G.H. Bower (Ed.), The psychology of learning and motivation (Vol. 26, pp. 125-173). New York: Academic Press.

Orne, M.T., Whitehouse, W.G., Dinges, D.F., & Orne, E.C. (1988). Reconstructing memory through hypnosis: Forensic and clinical implications. In H.M. Pettinati (Ed.), Hypnosis and memory (pp. 21-63). New York: Guilford Press.

Perlmutter, J., & Myers, J.L. (1973). A comparison of two procedures for testing multiple contrasts. Psychological Bulletin, 79, 181-184.

Petty, R.E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 19, pp. 123-205). New York: Academic Press.

Peynircioğlu, Z.F., & Tekcan, A. I. (1993). Revelation effect: Effort or priming does not create the sense of familiarity. Journal of Experimental Psychology: Human Learning & Memory, 19, 382-388.

Peynircioğlu, Z.F., & Watkins, M.J. (1988). Effect of unfolding stimulus presentation on recognition memory. In M.M. Grunberg, P.E. Morris, & R.Sykes (Eds.), Practical aspects of memory: Current research and issues (Vol. 2, pp. 518-523). Chichester, Sussex, England: Wiley.

Pirolli, P.L., & Mitterer, J.O. (1984). The effect of leading questions on prior memory: Evidence for the coexistence of inconsistent memory traces. Canadian Journal of Psychology, 38, 135-141.

Rabinowitz, J.C. (1989). Judgments of origin and generation effects: Comparisons between young and elderly adults. Psychology and Aging, 4, 259-268.

Rabinowitz, J.C. (1990). Effects of repetition of mental operations on memory for occurrence and origin. Memory & Cognition, 18, 72-82.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. Memory and Cognition, 21, 89-102.

Raye, C.L., & Johnson, M.K. (1980). Reality monitoring vs. discrimination between external sources. Bulletin of the Psychonomic Society, 15, 405-408.

Reason, J. (1990). Human error. New York: Cambridge University Press.

Reason, J., & Mycielska K. (1982). Absent minded. Englewood Cliffs, New Jersey: Prentice-Hall.

Schacter, D.L., Harbluk, J.L., & McLachlan, D.R. (1984). Retrieval without awareness: An experimental analysis of source amnesia. Journal of Verbal Learning and Verbal Behavior, 23, 593-611.

Schacter, D.L., Kaszniak, A.W., Kihlstrom, J.F., & Valdiserri, M. (1991). The relation between source memory and aging. Psychology and Aging, 6, 559-568.

Shaughnessy, J., & Mand, J. (1982). How permanent are memories for real life events? American Journal of Psychology, 95, 51-65.

Shimamura, A. (1989). Disorders of memory: The cognitive science perspective. In F. Boller & J. Grafman (Eds.), Handbook of neuropsychology (pp.35-73). Amsterdam: Elsevier.

Shimamura, A., & Squire, L. (1987). A neuropsychological study of fact memory and source amnesia. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 464-473.

Shimamura, A., & Squire, L. (1991). The relationship between fact and source memory: Findings from amnesic patients and normal subjects. Psychobiology, 19, 1-10.

Skinner, B.F. (1983). Intellectual self-management in old age. American Psychologist, 38, 239-244.

Slamecka, N., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 4, 592-604.

Snodgrass, J.G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. Journal of Experimental Psychology: General, 117, 34-50.

Sutton, C. (1984). A breakdown in symmetry. New Scientist, Jan. 26, 34-35.

Taylor, F.K. (1965). Cryptomnesia and plagiarism. British Journal of Psychiatry, 111, 1111-1118.

Tulving, E. (1983). Elements of episodic memory. Oxford: Oxford University Press.

Tulving, E. (1985). Memory and consciousness. Canadian Psychologist, 26, 1-12.

Tulving, E., & Thomson, D.M. (1971). Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 87, 116-124.

Warrington, E.K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. Neuropsychologia, 12, 419-428.

Watkins, M.J., & Peynircioğlu, Z.F. (1990). The revelation effect: When disguising test items induces recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 1012-1020.

Wells, G.L., & Loftus, E.F. (Eds.). (1984). Eyewitness testimony: Psychological perspectives. New York: Cambridge University Press.

Wells, G.L., & Turtle, J.W. (1987). Eyewitness testimony research: Current knowledge and emergent controversies. Canadian Journal of Behavioural Science, 19, 363-388.

Wells, G.L., Lindsay, R.C.L., & Ferguson, T.J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. Journal of Applied Psychology, 64, 440-448.

Winer, B.J. (1962). Statistical principles in experimental design. New York: McGraw-Hill.

Whittlesea, B.W.A., Jacoby, L.L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. Journal of Memory and Language, 29, 716-732.

Witherspoon, D., & Allan, L.G. (1985). The effect of a prior presentation on temporal judgments in a perceptual identification task. Memory and Cognition, 13, 101-111.