HOW CONTEXT CHANGES THE RETRIEVAL DYNAMICS OF A SOURCE

by

THOMAS WAYNE HANCOCK

(Under the direction of RICHARD MARSH)

ABSTRACT

Two experiments used a response signal methodology to explore the time-course of source-monitoring judgments. The basic thesis being investigated was that the revival rate of source information for a particular source would depend on the characteristics of a competing source. Both experiments tested a common source that was paired with either a dissimilar or similar competing source. Experiment 1 used a common perceptual source (i.e., seen) and revealed no advantage for additional revival time when paired with another perceptual source. However, source-monitoring performance did improve at later deadlines when the seen source was paired with an internal source. In Experiment 2, the source attributions for the common source (i.e., imagined) improved with additional time available for the revival of qualitative information, and did not depend on the competing source. The results suggest that the revival rate of source information is dependent on the context in which it is tested.

INDEX WORDS: Context effects, Source monitoring, Time-course, Memory

HOW CONTEXT CHANGES THE

RETRIEVAL DYNAMICS OF A SOURCE

by

Thomas Wayne Hancock

B.A., Saginaw Valley State University, 1994

M.A., Central Michigan University, 1997

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2002

HOW CONTEXT CHANGES THE

RETRIEVAL DYNAMICS OF A SOURCE

by

THOMAS WAYNE HANCOCK

Approved:

Major Professor:	Richard Marsh
Committee	Katherine Kipp Stuart Katz Zachary Estes Rob Mahan

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
May 2002

DEDICATION

To my parents, Larry & Carole Hancock

ACKNOWLEDGEMENTS

I would like to thank all of those who have directly or indirectly made it possible for me to obtain a PhD. To start, I would like to thank my major advisor, Richard L. Marsh, who has provided me with an invaluable educational experience. Second, I would like to thank the members of my doctorate committee for their valuable insight in critiquing this manuscript: Katherine Kipp, Stuart Katz, Rob Mahan, and Zachery Estes. Third, the companionship and intellectual input of my fellow graduate students has made my time at UGA, to say the least, entertaining. Those various graduate students are: Brian Jamieson, Marty Bink, Chris Mayhorn, Jenny Solberg, Jason Hicks, Scott Taylor, Jaime Durley, & Gabe Cook. Fourth, I would like to especially thank Carrie Noble for her invaluable support and friendship during my time here at UGA. Finally, to my family, Larry, Carole, Michael, & Angela who have provided unquestioning support.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Source monitoring recruits a set of flexible cognitive processes that are used in making attributions about the origin (or context) in which events were experienced (Johnson, Hashtroudi, & Lindsay, 1993). In the laboratory, source-monitoring experiments ask participants to make judgments concerning how a studied item was originally learned (i.e., to specify an attribute of the context in which it was experienced). For example, participants might be asked to recall whether stimuli were seen versus heard, spoken by a male versus a female, or presented in one location versus another. Failures of source-monitoring processes have been used to explain a variety of phenomena including assessment of the credibility and reliability of eyewitness testimony (Lindsay & Johnson, 1989), the genesis of inadvertent plagiarism (Marsh & Bower, 1993; Marsh & Landau, 1995; Marsh, Landau, & Hicks, 1997), distinguishing fact from fantasy (Johnson & Raye, 1981), and the occurrence of false memories (Hicks & Hancock, in press; Hicks & Marsh, 1999).

The source-monitoring framework suggests that people do not directly recall the specific source or origin of an event. Rather, to determine an event's origin, individuals evaluate the various qualitative characteristics associated with it. In this framework, source judgments are a form of attribution that are often based on incomplete details. The qualitative characteristics of memories that people assess include (but are not limited to) perceptual details, records of thoughts or reactions, emotions, and other details that were present when an event occurred. As such, two events from different origins should have unique characteristics associated with them in memory.

Differences in the amount and type of qualitative information that memories contain are used to make an attribution about whether an event occurred in one context versus another. The classic example of reality monitoring requires that participants distinguish between externally-generated (i.e., perceived) and internally-generated (i.e., mental) experiences. External events are generally associated with perceptual details (e.g., color, sound, texture, etc.) whereas internal events are associated with records of cognitive operations (e.g., imagination, elaborations, classifications, or interpretations). In order to engage in reality monitoring, an individual would have to retrieve from memory the qualitative characteristics associated with the original experience. If this qualitative information entailed vivid colors and sounds as opposed to cognitive operations, then the attribution that the event was perceived rather than imagined would be more likely to occur (Johnson, Raye, Foley, & Foley, 1981). In contrast, an attribution that the event was imagined would more likely be made if the qualitative information was comprised of more cognitive operations.

According to Johnson et al. (1993) it takes time to revive, or bring to mind, this source-specifying information. Therefore, two sources that contain different qualitative characteristics (e.g., external vs. internal) may revive (or be recollected) at different rates. The revival rate of source information is likely to depend on a number of different factors. The nature of the qualitative characteristics being retrieved will surely affect this rate. For example, emotional details may be retrieved more slowly than perceptual details, or perceptual details may be retrieved more slowly than semantic details. Moreover, research has indicated that different types of information can be retrieved at different rates. Jacoby and his colleagues have long argued that deep, recollective details revive more slowly than do less differentiated information that usually gives rise only to a feeling of familiarity (e.g., Jacoby, Toth, & Yonelinas, 1993; McElree, Dolan, & Jacoby, 1999).

Consequently, in the context of source monitoring, the revival rate of various qualitative characteristics should depend on how much recollective information a particular source contains. As such, measuring the revival rates of different sources would only represent a marginal extension of Jacoby's work. A more interesting possibility to investigate is that the identical source paired with two different candidate sources may revive at different rates. For example, a seen source that is paired with a heard source may revive at a different rate than when that same seen source is paired with a source in which items have been learned as anagrams that were solved at study. In other words, a perceptually seen source may revive at different rates depending on the other candidate source from which it is to be discriminated. One reason this might occur is because the amount of information that needs to accrue from memory may depend on the other source. An item presented in a seen context may need to accrue a maximal amount of information when paired with another perceived source such as heard items. The required amount of information that needs to be extracted may be less if the same seen item is paired with an internally generated source such as an anagram. In contrast, the alternative prediction is that the revival rate does not depend on the source-monitoring situation, but rather, is largely independent of it. Therefore, the purpose of the current study was to investigate whether or not different memorial characteristics associated with studied items revive at different rates and to what extent these effects are mediated by the particular combination of the sources to be discriminated.

The current research was motivated, in part, by a finding obtained by Marsh and Hicks (1998). In standard source-monitoring experiments, participants are usually given three response options for each test item. For example, if the two sources were seen and heard, then the test options would be seen, heard, and new. Marsh and Hicks found that if people were asked directly whether an item was seen (or heard) they had a yea-saying tendency to agree with whatever the query was for that test item. Of particular interest to the present investigation, they also found that accuracy was different for a seen source paired with a heard source than when the same seen source was paired with items studied as anagrams. Although this is an accuracy difference, the effect could manifest in revival rate differences as well. Thus, whether the revival rate for a particular source is dependent on the source from which it must be discriminated was the focus of the current study. However, it is possible that a given source may not behave differently in the context of two different sources with which it is paired. Whether this occurs may depend on the relative diagnosticities of the sources being considered. The relative diagnosticities are determined by how similar two sources are to one another, and this similarity is often thought to be related to which of the three basic types of source monitoring is being tested (e.g., Brown, Jones, & Davis, 1995).

The three types of source tasks include external-external discriminations where both sources are externally perceived (e.g., seen and heard), internal-internal discriminations where both sources reflect two different types of cognitive operations (e.g., judging items for how imageable they are or how categorizable they are), and internal-external discriminations where one source is internally derived and the other is externally derived (e.g., seen and imagined items). Whether a source behaves differently in the context of two different sources may depend on the type of source-monitoring task being tested. Therefore, across different experiments, the type of source-monitoring task being tested was varied to determine whether this factor contributed to differences in the relative revival rates of source information.

Based on the foregoing analysis, the time required to revive associated characteristics could be determined by the combination of sources used. Sources that share similar characteristics may have few diagnostic details to differentiate them, and this may result in slower source-monitoring reaction times for both origins. This outcome might occur if the amount of time required to make a source judgment is positively correlated with the necessary amount of information that needs to be retrieved in order to make the discrimination. When sources share similar characteristics it may require a greater amount of time to accumulate enough information to distinguish between them. However, when the combination of sources are diagnostically quite different (e.g., internal vs. external sources) then differences in revival rates should become apparent earlier in the time course of making a source judgment because less information is needed from memory before a source decision can be made.

One method for investigating the revival rates of different studied sources is the response-signal paradigm (Reed, 1973). After the study phase of a response-signal experiment, each stimulus is presented and participants are instructed to respond immediately with their decision when prompted. The interval of time between item presentation and the onset of the prompt to respond is systematically varied from trial to trial. This procedure allows researchers to examine accuracy as a function of different response signal deadlines (i.e., the parametric manipulation of the time available to make the decision). Theoretically, shorter deadlines should allow for less memorial information to be retrieved (revived), and therefore, judgments are based on less information retrieved from memory. As the deadline is parametrically increased, greater amounts of information can be retrieved which should increase the quality of the information on which the speeded judgments are made. In the limit case (i.e., longest deadline), performance should asymptote to a level of performance that would have been observed had no deadlines been imposed at all.

This methodology has been used in only a couple of investigations of the revival rate of qualitative characteristics (Hintzman & Caulton, 1997; Johnson, Kounois, & Reeder, 1994). In Hintzman and Caulton's (1997) study, participants received several study blocks where they studied either auditory or visual word stimuli. During test, the response-signal paradigm was applied either to a recognition test or to a source-monitoring test. They found that above-chance performance on the recognition test was obtained at shorter deadlines than on the source-monitoring test. These results imply that recognition judgments and source-monitoring judgments do not necessarily rely on the same type or quantity of information. As Johnson et al. (1993) claimed, recognition performance depends on less differentiated input than does source performance. Because less information is needed to render an old-new recognition judgment, the recognition decision can be completed earlier than the source decision which requires more differentiated information in order to respond. However, this study also demonstrated that accuracy measures on the two studied sources (seen and heard) did not differ as a function of the response signal. In other words, seen and heard information appeared to revive at similar rates.

Johnson, Kounios, and Reeder (1994) also investigated the revival rates of different characteristics in a source-monitoring study. During study, participants either perceived a set of pictures or imagined them from word labels. At test, they were required to respond "Perceived," "Imagined," or "New." They found that imagined studied items revived more quickly because performance was better at shorter response-signal deadlines than it was for perceived items. As noted before, this study illustrates that different qualitative characteristics do not necessarily revive at the same rate.

Both the Hintzman and Caulton (1997) and Johnson et al. (1994) studies contained some odd infelicities. For example, in the former case, the auditory source and visual source were blocked at encoding. Therefore, in addition to perceptual details that could be used on the response-signal test, temporal information for the source (which was learned first vs. second) could also be affecting the results. Likewise, Johnson et al. compared various deadlines but analyzed their data not on the proportion correct at each deadline, but rather, on transformed data from multinomial models of source discrimination which requires making various assumptions that might be questionable (Batchhelder & Riefer, 1990; Kinchla, 1994). The important point is that these derived scores are one step removed from the raw data.

EXPERIMENT ONE

The first experiment had two objectives: (1) to replicate conceptually Johnson et al.'s (1994) findings which demonstrated a quicker revival rate for internally generated items compared to externally perceived items, and (2) to assess whether the revival rate of a particular source depends on the source with which it is paired. The latter objective is the novel and more important of the two objectives.

Materials and Methods

Participants. Seventy-five undergraduate participants (38 in the seen-heard condition and 37 in the seen-anagram condition) from the University of Georgia participated in exchange for partial credit towards a course research requirement. Participants were tested individually in sessions that lasted approximately 40 min.

Materials and Design. The design for this experiment was a 2 (source combination: anagram-seen vs. seen-heard) X 4 (response signal lag: 200, 600, 1000, 1400 ms) mixed factorial with source combination as the between-subjects variable and response signal lag as the within-subjects variable. One hundred eighty medium frequency words were chosen from the Kucera and Francis (1967) corpus. Participants received 60 items randomly assigned to each source at study; randomization was repeated so that each participant received a unique combination of items. For the source-monitoring test all 120 studied items plus an additional 60 distractor items were randomized anew and presented individually. Thus, each participant received 15 test items at each of the four deadlines from each of the three sources. For any given participant, lag was randomly assigned across trials by the software.

Procedure. Participants were informed that they were going to be participating in a memory experiment that consisted of two phases. The first phase consisted of a practice phase to acquaint them with the task that they would be performing during the experiment. Participants initially studied 24 words at a 3 second presentation rate. In the external-internal condition, half of these items were seen intact whereas the remaining items appeared with two letters interchanged (i.e., anagrams). These two letters were denoted by a caret (^) underneath them. Participants were informed to interchange the two letters to determine the word's identity, and once it was determined, say the word to the experimenter and then press the space bar to continue on to the next trial. In the external-external condition the combination of sources were seen intact items and heard items spoken by the experimenter. All items in the conditions were preceded by a short beep sound and fixation point for 250 ms. This procedure helped to ensure that participants paid attention to the study sequence.

The practice test phase occurred immediately after study and the practice test items consisted of the original 24 studied words along with 12 new words. Participants were told that they were going to see a sequence of words, some that were previously studied and others that were not. For each of the words they were asked to make a source judgment. For instance, in the external-internal condition the task was to decide whether the word was previously seen intact, generated from an anagram, or was a new item. Labels were placed on the "J", "F", and spacebar keys, respectively. They were told to place their left index finger on the "F" key, their right index finger on the "J" key, and both thumbs on the spacebar. Participants were further instructed to withhold their response until they heard an auditory tone, but to respond immediately on hearing the tone. In order to ensure prompt responding, participants taking longer than 350 ms after the signal received the feedback “response too slow” (cf., Hintzman & Caulton, 1997). After completion of the practice test, any procedural questions from the participant were answered and then the actual experiment began. At the end of the experiment the participants were debriefed and thanked.

Results and Discussion

Trimming was performed on the data in order to reduce the effects of participants who, on average, took more than the allowed processing time for each response signal deadline. For each participant, the mean and standard deviation was calculated for each response signal deadline. Three participants in each condition, whose mean RT exceeded 500 ms for any signal response deadline, were removed from the analysis (cf. Hintzman, Caulton, & Curran, 1994). After trimming, 34 participants remained in the seen-anagram condition and 35 participants remained in the seen-heard condition.

Another concern was determining the appropriate measure of source-monitoring performance. Typically, measures of source-monitoring performance divide the number of correctly identified items (e.g., seen studied items identified as seen during test) by the total number of items in that class of items (e.g., seen) that were originally studied (Marsh & Hicks, 1998). However, this measure is inappropriate for the current methodology because old-new recognition performance varies greatly across the various deadlines. Specifically, at the shortest response deadlines fewer items are recognized as old compared with longer deadlines that approach asymptotic performance (Johnson, Kounios, & Reeder, 1994). As a result, standard source-monitoring scores at the earlier response signal deadlines will have fewer items identified as old and, thus, will appear to have inferior source accuracy as compared with later deadlines. To correct for this, single-source conditional source identification measures (CSIM) were used, because these scores are independent of recognition performance across the various deadlines (Murnane & Bayen, 1996). For each deadline, the CSIM score for each source represents the number of items correctly identified (e.g., seen items labeled as seen) divided by the number of items identified as old (e.g., seen studied items labeled regardless of the source specified).

In order to obtain additional stability in the interpretation of information accrual over time the data from two adjacent response signals were averaged (Roediger & McDermott, 1995; Tufte, 1983). This practice is often used when a continuous underlying dimension (in this case time for stimulus onset) is necessarily represented by several arbitrarily chosen discrete values (in this case four within-subject deadlines). For example, the data from the first and second deadline were averaged for each participant, the same for the second and third, and so forth. Therefore, the following statistical analyses were conducted on the three resulting averages instead of the original four deadlines. The results and discussion section begins with the critical comparisons of central interest, follows with the individual condition analyses, and concludes with the theoretical discussion.

Critical Tests. The specified a priori critical comparisons testing the novel hypothesis of this study were analyzed with a 2 (condition) X 3 (signal) mixed model ANOVA with condition as a between-subjects variable and response signal deadline as a within-subjects variable for the seen sources that were otherwise identical in both conditions (see Figure 1). A main effect of condition was found, $F(1,67) = 4.04$, with seen items studied in the context of anagrams (seen-anagrams) (.72) having better source accuracy than seen items in the context of heard (seen-heard) (.66). This result is consistent with Marsh and Hicks's (1998) finding that different levels of source-monitoring performance are dependent on the particular combinations of sources tested together. In addition, there was a main effect of response signal indicating that participants were able to accrue additional information as a function of the lag between the presentation of the stimulus and the response signal, $F(2, 134) = 8.03$. Further analysis revealed an improvement in source accuracy between the first (.66) response signal and the second (.71), $t(68) = 3.32$. There was no difference between the second response signal and the third (.71), $t(68) = 0.296$.

However, the main effect of lag is qualified by a significant interaction between condition and response signal deadline, $F(2, 134) = 6.19$. Subsequent analysis revealed a significant linear, $F(1, 33) = 16.102$, and quadratic trend, $F(1, 33) = 8.76$, for seen items learned in the context of anagrams whereas performance for seen items learned in the context of heard items did not significantly change across deadlines. Additional analyses revealed no significant differences in performance between the two seen sources at the first response signal, $t(68) = 0.186$. However, the seen source when combined with anagram items had better source accuracy than when it was paired with heard items in both the second and third deadlines, $t(68) = 3.013$ and $t(68) = 2.465$ respectively. Therefore, source accuracy for seen items in the seen-anagram condition improved with the retrieval of additional contextual information obtained from the longer response signal deadlines. This was not the same pattern of results found for the seen items in the seen-heard condition. In that condition, the competing source (i.e., heard items) possibly shared too many similar, perceptual characteristics with the seen items, and therefore, the amount of information accrued at the longer deadlines was unable to further distinguish between the two sources.

Condition Analyses. Having described the critical comparisons, the statistical analysis for each condition follows. Considering each condition in turn, the seen-anagram condition that is summarized in Figure 2a was analyzed with a 2 (source) X 3 (signal) within-subjects ANOVA that revealed no effect of source, $F(1,32) = 0.05$, but a main effect of response signal, $F(2, 66) = 17.79$. Analysis of the signal response deadlines revealed an improvement in source accuracy from the first response signal (.67) to the second signal (.75), $t(32) = 5.72$, as well as the third (.75), $t(32) = 4.65$. Once again, there were no significant differences between the second and third response signals, $t(32) = 0.266$, suggesting that performance had asymptoted prior to the longest lags.

The interaction between source and response signal approached significance in this condition, $F(2,66) = 3.11$, $p = 0.051$. Further analysis revealed that over the deadlines for the seen items there was a significant linear trend, $F(1,32) = 16.1$, as well as a quadratic trend, $F(1,32) = 8.76$. Improvement in source accuracy occurred between the first signal response deadline (.65) and the second (.76), $t(32) = 5.09$, as well as the third (.75), $t(32) = 4.01$; however, no difference between the second and third deadlines were found, $t(32) = 0.10$. These results that were obtained for the seen items suggest that a significant amount of contextual information was revived from memory during the interval of time between the first and second response signals. However, after the second deadline additional time did not further enhance source accuracy. For the anagram items there was also a significant linear trend, $F(1, 32) = 7.22$, but the quadratic function was not significant. Similar to the results for the seen source, results for the anagrams indicated there was an improvement in source accuracy between the first signal response deadline (.69) compared to the second (.73), $t(33) = 2.88$, and third signal (.74), $t(33) = 2.69$. However, there were no significant changes in source accuracy between the second and third response signals, $t(34) = 0.27$.

For the seen-heard condition summarized in Figure 2b, a 2 (source) X 3 (signal) within ANOVA revealed no effect of source, $F(1, 33) = 0.19$. This pattern of performance was predicted based on Hintzman and Caulton's (1997) study that found equivalent revival rates for both seen and heard sources. Furthermore, there was a main effect of response signal, $F(2, 68) = 11.34$. Subsequent analysis revealed a significant improvement in source accuracy across all three deadlines, between the first signal (.63) and the second signal (.67), $t(33) = 3.35$, the first and third signal (.69), $t(33) = 3.75$, and between the second and third signal, $t(33) = 1.99$. A significant interaction was found between the source and the signal response deadline, $F(2, 68) = 5.45$. Analysis of the seen items as a function of response signal revealed no significant differences in source accuracy across the response deadlines, $F(1, 33) = 0.19$. However, for the heard items there was both a linear trend across response deadlines, $F(1, 33) = 16.95$, as well as a quadratic trend, $F(1, 33) = 5.26$. Further analysis revealed a difference between the first deadline (.60) and the second (.69), $t(33) = 4.48$, and between the first and third (.72), $t(33) = 4.12$. A modest difference was also found between the second and third deadlines, but it did not reach conventional levels of significance, $t(33) = 1.71$, $p < .09$. The pattern of results for the seen-heard sources revealed a consistent level of performance for the seen items across the deadlines, while source accuracy for the heard items steadily improved across the deadlines.

Summary. Participants had different source-monitoring accuracy for the seen source depending on the characteristics of the other source with which it was paired during study. Source accuracy for an external source (i.e., seen items) studied in the context of a highly diagnostic source (e.g., an internal source) was initially rather poor when retrieval time was limited. However, with additional time participants were able to retrieve adequate diagnostic information to discriminate the external source from the internal source. In the external-external source condition, participants had difficulty identifying the seen source from another external source (i.e., heard items). These results suggest that source-monitoring accuracy for an external source is benefited by additional retrieval time in the context of a highly diagnostic, internal source. Although, the same benefit of additional information at the longer response deadlines does not occur in the external-external source context.

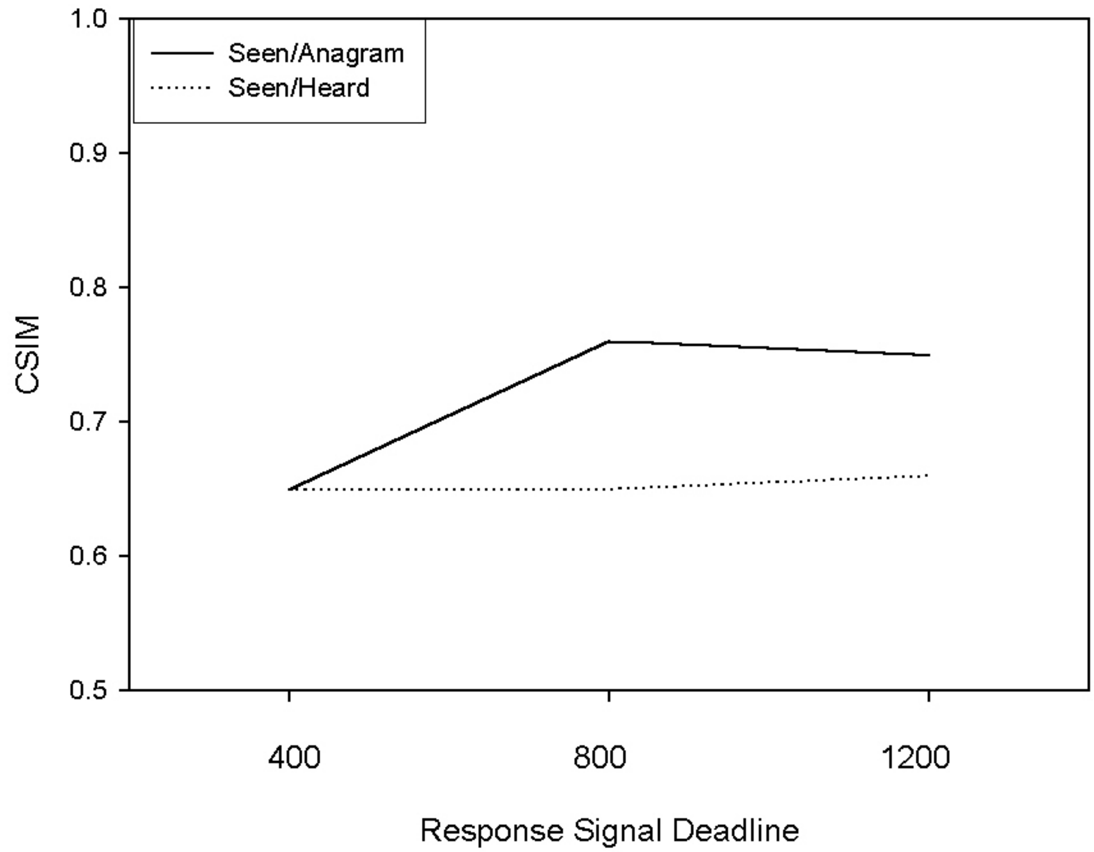
There are a number of ways to interpret this outcome. First, in the seen-heard condition, after identifying an item as old, all participants have to examine is perceptual information. If perceptual information, regardless of whether it is seen or heard, is shared in similar formats or in similar ways, examination of that information and discrimination between the sources may be facilitated. As such, additional time may be of little further utility. In contrast, if the seen-anagram information is stored in different formats, perhaps it is more difficult to compare. If this were so, additional time would be beneficial. Thus, the results of this experiment could be a consequence of how information is represented in memory. Second, the difference in diagnosticity of the seen-anagram condition could cause participants to engage in a serial search for information starting with the most diagnostic source (i.e., anagrams). Failing to find evidence of cognitive operations, participants may then look for the weaker evidence (i.e., seen items) (for similar arguments see Feller, 1961; Johnson & Raye, 1981; Marsh & Hicks, 1998). If such a serial search were taking place, it would explain why seen items benefit from additional memorial search time.

Continuing in this fashion, in the seen-anagram condition at the earliest deadline, participants are inspecting their memory for any evidence of cognitive operations, which have been previously shown to revive quicker than perceptual information (Johnson, Kounios, & Reeder, 1994). Upon failure to find any evidence of cognitive operations, participants then begin to inspect their memory for perceptual information. Therefore, for seen items it is not until the later response signal deadlines, after the participants fail to find evidence of cognitive operations, that source accuracy begins to improve for the seen source. However, this is not the case for the two perceptual sources. In this instance, both sources access similar information, with neither source having privileged access over the other competing source (Hintzman & Caulton, 1997). Because neither of the perceptual sources can revive from memory unique information that is diagnostic of a specific source, participants are unable to easily distinguish between the two external sources.

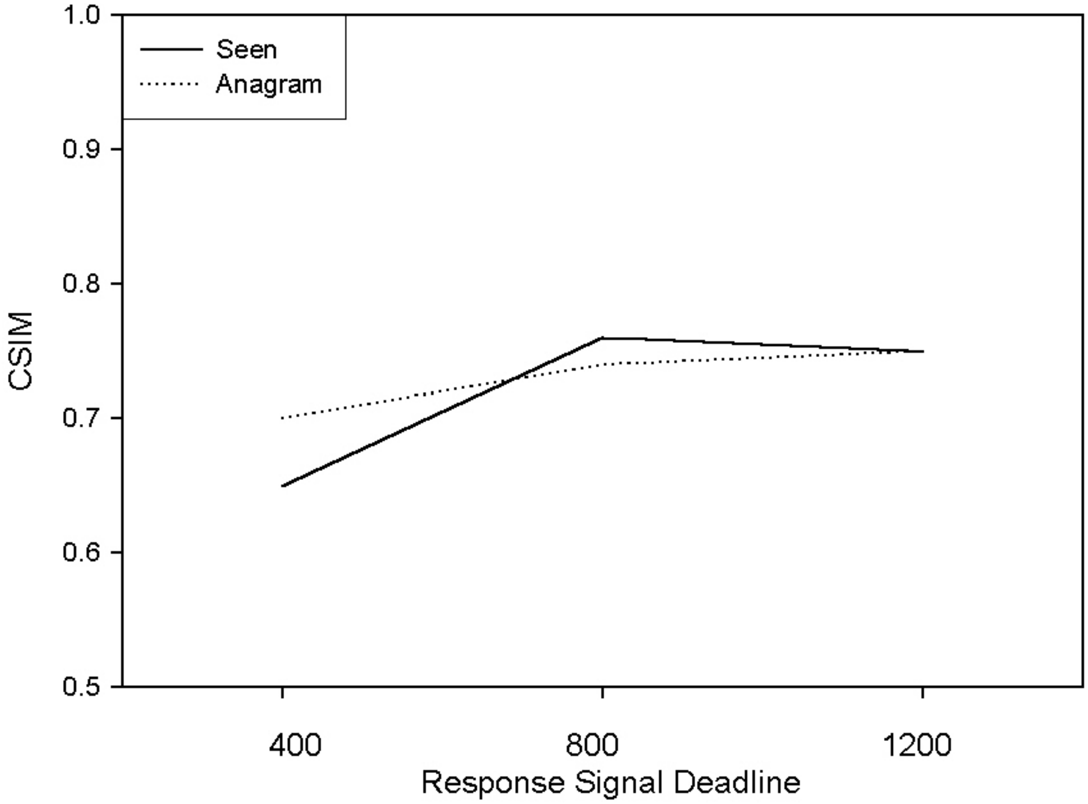
An alternative interpretation is that information was not reviving at different rates, but rather, the time course differences were a result of the source attribution decision process. For instance, in the seen-heard condition, by the first response deadline participants may have revived all the information that was available in memory.

Therefore, the subsequent deadlines did not represent the revival of information, but instead the difficulty in distinguishing between the two sources. The two competing perceptual sources shared several overlapping characteristics, therefore, the process of making a source attribution was difficult. At test, performance was not further facilitated by additional time because participants lacked enough diagnostic information on which to base their decisions. Similarly, in the seen-anagram condition, participants may have revived all of the diagnostic source information by the first deadline, and it was not until approximately between second and third deadlines that they were able to make correct source attributions. Regardless of whether the results favor revival rates or decision processes, Experiment 1 clearly demonstrates that correct source attributions for a seen perceptual source is influenced by the specific characteristics associated with a competing source.

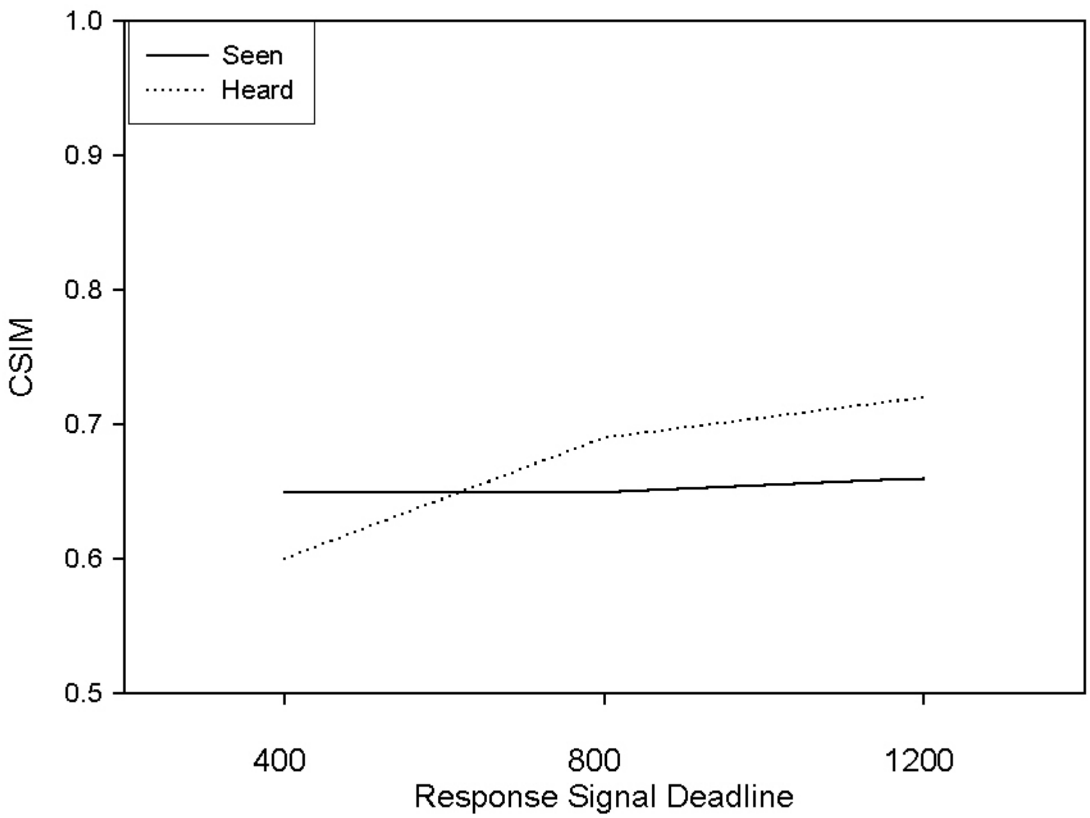
Experiment 1
Seen Items
CSIM



Experiment 1: Seen/Anagram
CSIM



Experiment 1: Seen/Heard
CSIM



EXPERIMENT TWO

The purpose of Experiment 2 was to explore whether the critical outcome from the first experiment would generalize to different source-monitoring conditions. In the first experiment, the revival rate of an external source (i.e., seen items) was examined in two different contexts (i.e., generated vs. heard items). In Experiment 2, the common source was one that relies primarily on internally-generated information. Participants were asked to imagine pictures from their word label referents (Johnson, Raye, Foley, & Kim, 1982). In one condition, this internal source was paired with actually seeing pictures (external) in an internal-external source-monitoring condition and with a frequency judgment task in an internal-internal source test. Based on the results from Experiment 1, the revival rate of the imagined items in the context of the other internal source was predicted to be slower than when the alternative was an external source. In the external-external condition in Experiment 1, source accuracy did not improve across deadlines as a result of overlapping features, therefore a similar prediction could be made to an internal-internal condition in Experiment 2, because both sources will rely on similar characteristics to make source attribution claims. Thus, this experiment used the general logic of examining the revival rate of a common source (imagination) between two different combinations of sources with the prediction that the imagined source would behave differently when paired with the internal frequency judgment source compared with the external picture source.

Materials and Methods

Participants. Seventy-six undergraduate participants (38 per condition) from the University of Georgia participated in exchange for partial credit towards a course research requirement. Participants were tested individually in sessions that lasted approximately 40 minutes.

Materials and Design. The basic design replicated Experiment 1 with the exception of the source combinations (i.e., seen-imagined vs. frequency-imagined). One hundred-eighty items were chosen from the Snodgrass and Vanderwart (1980) norms. These norms contain labels for common objects (e.g., bicycle) along with black and white line drawings of the items. Three sets of 60 items were equated on the variables established in these picture norms including picture complexity, imageability, and frequency of occurrence. Two of the three sets were used during the study phase while the third was used as new items during test (i.e., distractors).

Procedure. As in Experiment 1 a practice phase preceded the experimental phase. Prior to the practice study list, participants were instructed that they would be looking at pictures on some trials and imagining pictures from word-labels on other trials. At that time, participants were provided with two examples of what the word labels and corresponding pictures would look like from the Snodgrass and Vanderwart (1980) norms. After being shown the examples, participants were informed that imagined pictures in the imagery condition should be line drawings similar to the ones provided by the examples. Finally, participants were asked to rate how long it would take an artist to draw the picture (or image) (Johnson et al., 1982). Therefore, on every trial during the 6 s after the word label had been removed, participants would estimate the time it would take to draw the image using the following scale: 1 (very quick) to 5 (very long) with the range of numbers representing intermediate levels of time.

For the seen-imagery condition, participants were presented with 24 single-word object labels during the practice study phase. On half of these trials, participants saw the corresponding picture subsequent to reading the word label. On the other half of the trials, a picture did not follow the word label, rather a blank screen signaled the participant to imagine a picture corresponding to the word label. On each trial the word label was presented for 2 s followed by a 6 s display that was either a picture or a blank screen.

The practice test phase immediately followed the practice study phase and closely followed the procedure from Experiment 1. Participants were tested only with the word labels and not the corresponding pictures. The practice test list was comprised of all of the 24 previously studied word labels along with 12 new distractor words. Participants were instructed to judge whether each word was previously seen as a picture, imagined, or was a new item. The same four response deadlines were used as in Experiment 1 (i.e., 200, 600, 1000, 1400 ms). After the practice phase was completed, participant questions were answered and the study and test phases commenced.

For the frequency-imagery condition, the study phase was similar to the seen-imagery condition with the exception that a screen without any pictures always appeared after the word labels for 6 s. For example, during both the imagery trials and frequency judgment trials, the word labels were followed by a screen instructing participants to perform the appropriate task. The imagery instructions replicated the imagery procedure used in the seen-imagery condition. For judgments of frequency, participants were asked how often they encountered the object in the last 2 weeks. Their response options ranged on a Likert scale from 1 (very infrequent) to 5 (very frequent), with 3 being moderate. The screen following each word label would display either the word 'IMAGERY' or 'FREQUENCY' in the center of the screen to inform the participant of the appropriate judgment required for the item. During test, participants evaluated each word with regard to whether it had been previously rated for frequency, had been imagined, or was new.

Results and Discussion

For each participant the mean and standard deviation was calculated for each condition. Using the same trimming procedures described in Experiment 1, two participants in the imagery-seen condition and three in the imagery-frequency condition, whose mean RT exceeded 500 ms, where excluded from further analysis leaving a sample size of 71 (36 in the imagery-seen condition and 35 in the imagery-frequency condition).

Critical Tests. The a priori critical comparison was between the common source (i.e., imagined items) as a function of the two different study conditions. A 2 (condition) X 3 (signal) mixed ANOVA with condition as a between-subjects variable and signal as a within-subjects variable was used to analyze the imagined source data (see Figure 3). A significant interaction was not found between the condition and response signal deadlines, $F(2, 138) = 0.59$. However, a main effect of condition was found, $F(1, 69) = 4.79$, with imagined items studied in the context of seen (.74) having better source-monitoring accuracy than imagined items in the context of frequency judgments (.66). A main effect of signal was also evident, $F(2, 138) = 20.51$. Further analysis revealed a significant improvement in source accuracy across all three signal response deadlines, from the first response signal (.65) to the second (.70), $t(70) = 4.16$, and from the second to the third (.74), $t(70) = 3.52$. Subsequent analysis revealed a significant linear trend for the imagined source in both the imagined-seen condition, $F(1, 35) = 17.855$, and the imagined-frequency condition, $F(1, 34) = 8.982$. However, the quadratic trend for both conditions did not approach significance.

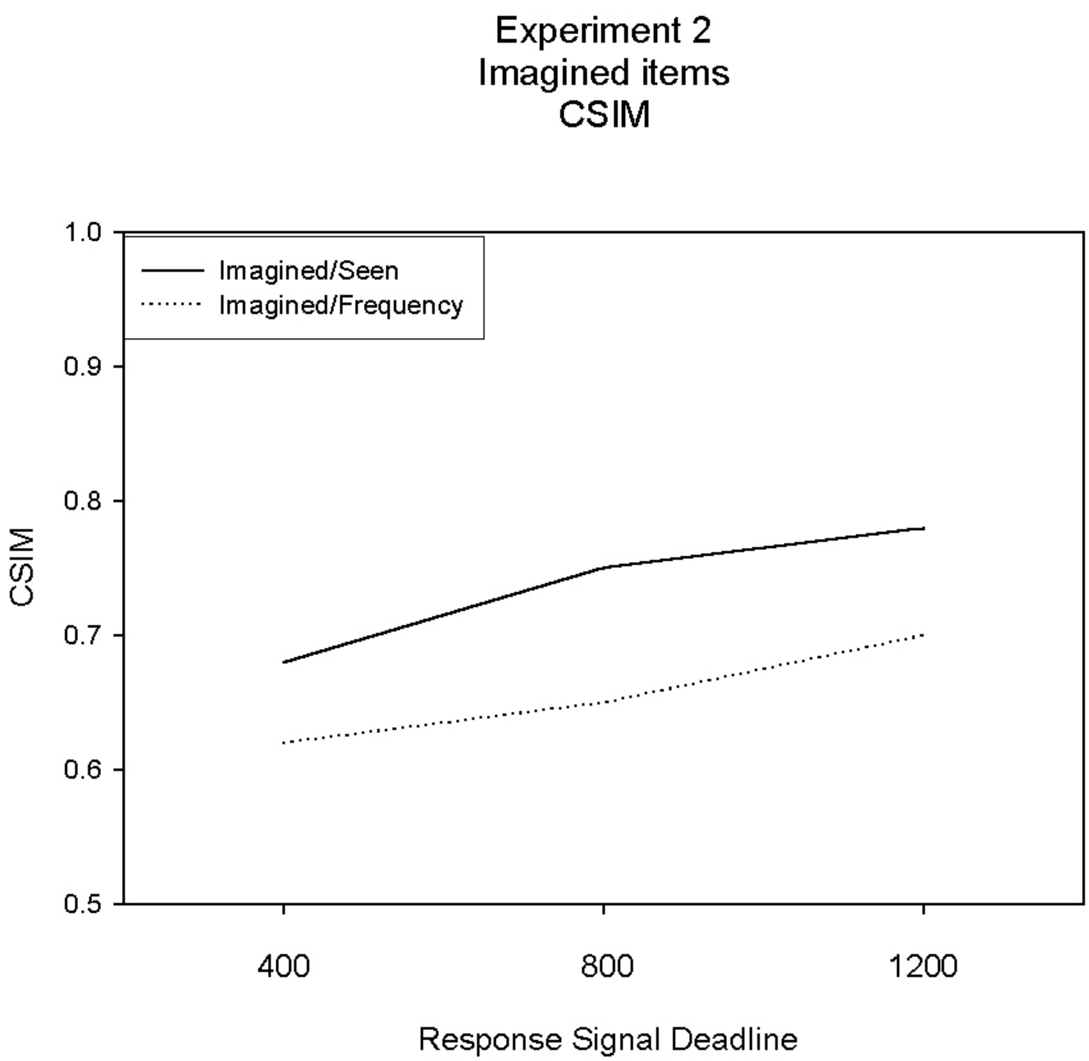
This general pattern of results replicates the first experiment by demonstrating that the revival rate of internal contextual information for a specific source (e.g., imagined items) has privileged memorial access when paired with a dissimilar external source (e.g., seen items). In addition, the main effect of response signal and linear trends combined with the lack of a significant interaction between the imagined items in the two conditions suggests that the common imagined information revived at approximately the same rate in the two conditions. However, imagined items in the imagined-frequency condition were not easily distinguishable from one another, thus, there was overall lower accuracy in this condition.

Condition Analyses. Considering each condition in turn, a 2 (source) X 3 (signal) within-subjects ANOVA for the imagined-frequency condition summarized in Figure 4a revealed no interaction between the studied source and the signal response deadline, $F(2, 68) = 0.25$ along with no effect for source, $F(1, 34) = 1.97$. However there was a main effect of response signal, $F(2, 68) = 23.89$. Subsequent analysis revealed a difference between all three deadlines, with a consistent increase from the first signal (.58) to the second signal (.63), $t(34) = 4.20$, and also from the second to the third signal (.67), $t(34) = 3.94$. This pattern of results suggests that the two internal sources were not easily distinguishable from one another, however, as the information accrued across the response signals source accuracy improved.

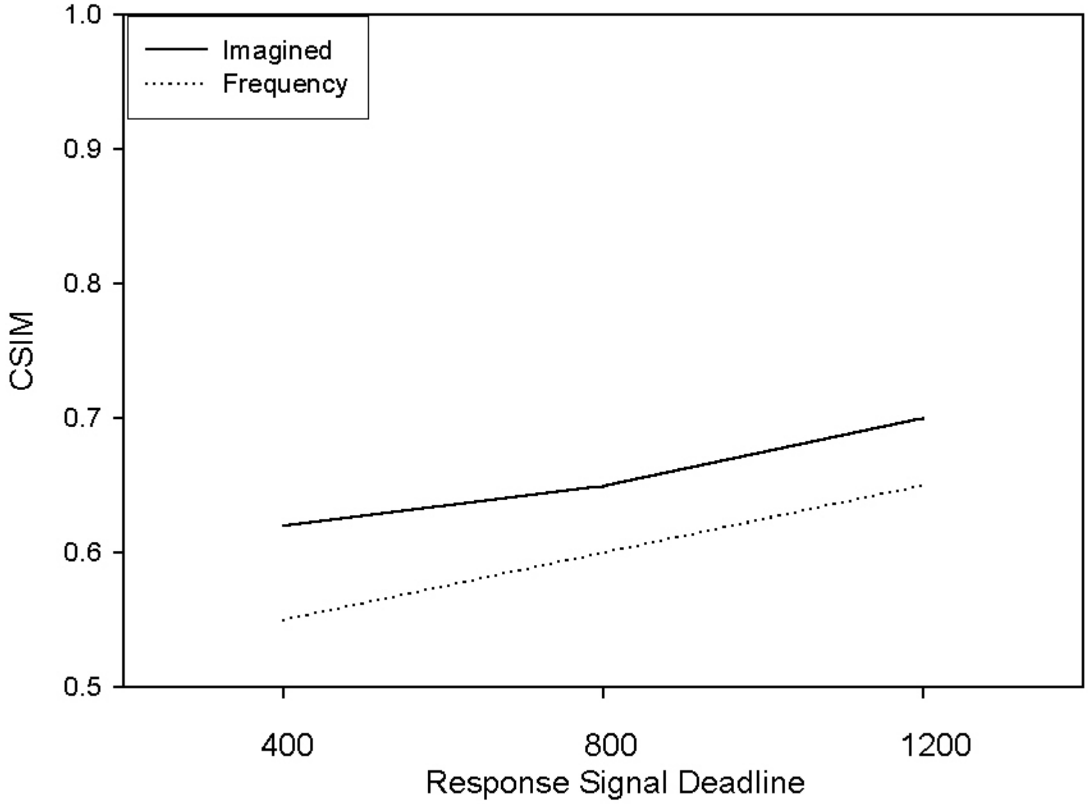
The imagined-seen condition was analyzed by a 2 (source) X 3 (signal) within-subjects ANOVA. The analysis did not reveal a significant interaction between source and signal response, $F(2, 70) = 1.91$. However, there was an effect of source, $F(1, 35) = 7.4$, with studied imagined items (.70) having better source-monitoring performance than seen items (.62) (see Figure 4b). There was also a main effect of response signal deadline, $F(2, 70) = 32.08$. Additional analysis revealed a significant increase in performance across all three deadlines. Performance at the first response signal (.64) improved at the second signal (.69), $t(35) = 4.2$, and increased more so between the second and third response signal (.73), $t(35) = 5.27$.

Summary. These findings demonstrate that when a common internal source (i.e., imagined items) is paired with an external source (i.e., seen items) it is able to accrue more diagnostic, or distinguishing, information quicker than when the same source is paired with an internal source. When the common internal source is studied with another internal source the similarity, or overlapping characteristics between the imagined items and frequency judgment items attenuated performance. Therefore, neither source had an advantage in accessing contextual information from memory compared to its competing source.

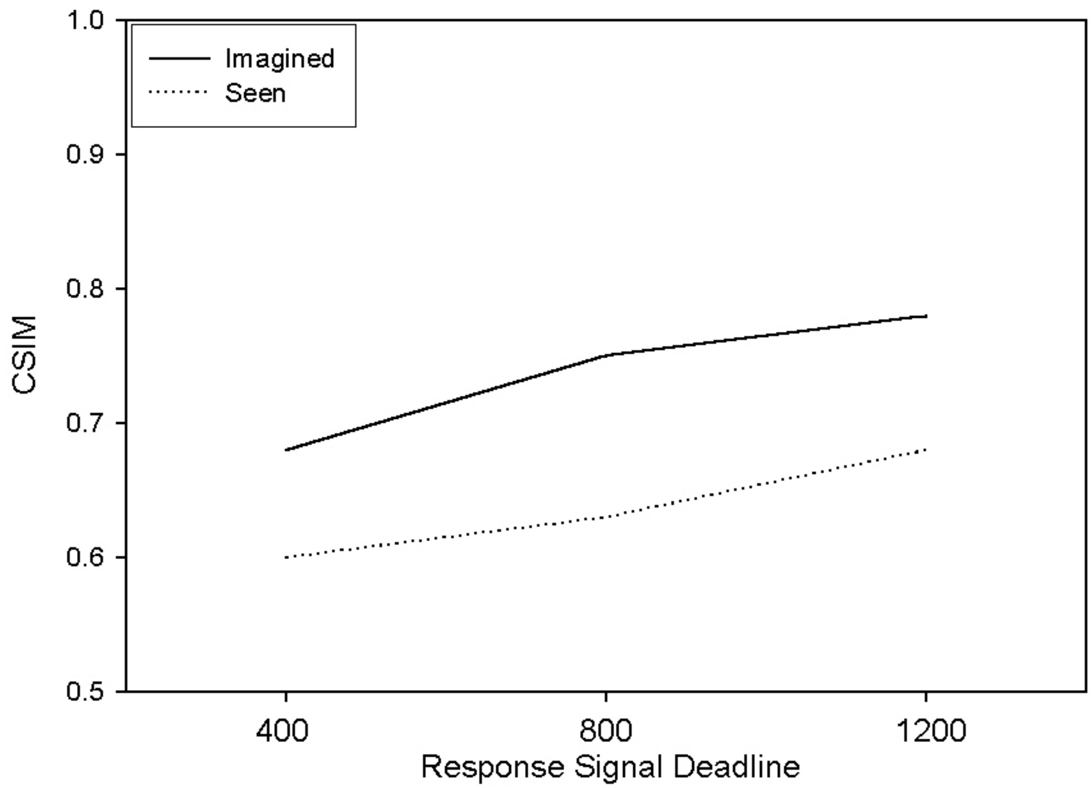
The lack of a significant interaction between the imagined sources in the two conditions further illustrates this point. The main effect of source accuracy for the imagined-seen condition demonstrates that it was a relatively easy process to distinguish between the two sources across all the deadlines (i.e., there are intercept differences and not slope differences). However, in the imagined-frequency condition the process of making source judgments between the two competing items was more difficult because both sources contained cognitive operations. As a result, this difficulty in distinguishing between the two sources could have been a factor in lowering source accuracy performance at the earlier deadlines. At the earliest response signals participants were unable to accrue enough diagnostic, or unique, information associated with a specific source. Therefore source-monitoring performance was low; however, when provided with additional retrieval time participants were able to revive more diagnostic information specific to an individual source and thereby improve overall performance.


Experiment 2
Imagined items
CSIM
Imagined/Seen
Imagined/Frequency
CSIM
1.0
0.9
0.8
0.7
0.6
0.5
400
800
1200
Response Signal Deadline

Experiment 2: Imagined/Frequency
CSIM



Experiment 2: Imagine/Seen
CSIM



GENERAL DISCUSION

In each experiment a common source was presented to participants in one of two different contexts. In one context the other source was easily distinguishable from the first, while in the second context the two sources shared some overlapping qualities that rendered the sources less easily distinguished from one another. At test, the amount of time allotted to make a source attribution was manipulated. The research goal was to investigate the retrieval dynamics of a particular source when the characteristics of the competing source were altered. These two experiments not only revealed the effects of retrieval constraints on the process of making source attributions, but also demonstrated the influence of contextual information associated with the competing source.

During Experiment 1 a common perceptual source (i.e., seen) was paired with another perceptual source (i.e., heard items) or with an internal source (i.e., anagrams). In the external-external case, source-monitoring performance did not fluctuate as a function of response signal deadline. However, when the perceptual source was paired with an internal source a different pattern arose. In the external-internal case, correct performance improved as the deadline was increased.

In Experiment 2, the revival rates of a common internal source (i.e., imagined) were examined as a function of the competing source with which it was paired. Unlike Experiment 1, there was no interaction between the common internal sources across the two conditions. Rather, the internal source in the imagined-seen condition had significantly better source-monitoring performance across all response deadlines compared to the imagined items in the imagined-frequency condition. In addition, the source accuracy for the common sources in Experiment 2 consistently improved across all three deadlines. Thus, the pattern of results for the common source in Experiment 2 was different from Experiment 1 in several aspects.

The current study demonstrates that any given source does not have a pre-specified "amount" of contextual information that is necessary to make a correct source attribution. If this were so, then the common source in both experiments should have displayed similar performance in both conditions as a function of changing the response deadlines. In other words, source attribution claims would have been independent of the other source with which it was paired. However, such an outcome was not obtained.

When the studied characteristics of two sources varied greatly from one another (e.g., perceived and imagined) source judgments were relatively easy to perform. The current results suggest that these judgments can be made with relatively less information accrued from memory, because less information is needed to discriminate between the two sources. For example, if the revived qualitative information entailed strong cognitive operations as opposed to vivid details and sounds, one would be more likely to determine that the event was imagined rather than perceived (Johnson, Raye, Foley, & Foley, 1981). This was the case for the internal source in the context of an external source in Experiment 2. In this condition, the overall source accuracy was significantly greater compared to when the common internal source shared similar characteristics with the competing source (e.g., imagined-frequency condition).

However, the ease of distinguishing between sources does not always result in greater source accuracy judgments. In the external-internal condition of Experiment 1, it was not until the later response deadlines that source accuracy performance for seen items began to improve. The pattern of results suggest that it was not the discriminability of sources that influenced performance, for if this were the case then both the internal and external sources should have had equivalent source accuracy across all the deadlines. However, the current results are consistent with the idea that participants originally inspected their memory for evidence of cognitive operations. It was not until the failure to find any significantly diagnostic evidence of cognitive operations that participants began inspecting their memory for other details, such as perceptual information, which led to improved source accuracy for seen items at the later deadlines.

When two sources share similar characteristics, the difficulty of making source-monitoring judgments (e.g., internal-internal condition) should increase. In these instances, participants require more information to determine an event's origin. Therefore, this requirement should directly affect the amount of accrued information necessary to perform a source-monitoring judgment. Stated another way, revival rates should be positively correlated with differences in the difficulty to make source judgments. When source attribution judgments increase in difficulty this should translate into a greater amount of time necessary to revive enough information to distinguish between the sources.

The data from the present experiments partially support this argument. When two perceptual sources (seen-heard) were tested in Experiment 1, source accuracy for the seen source never improved with the additional processing time provided by the later deadlines. This outcome suggests that these two sources have a significant number of shared qualities between them. In contrast, the common imagined source in the internal-internal condition of Experiment 2 showed lower source-monitoring performance compared to the imagined items in the internal-external condition. However, at the later deadlines, source accuracy did improve, indicating that participants were able to revive more diagnostic information. Therefore, unlike the external-external condition that was not able to revive additional diagnostic information of a specific source with additional retrieval time, it was possible in the internal–internal condition. This might not be the case for all internal sources, but, it seems that imagery and frequency judgments are sufficiently different from one another such that additional processing time does improve participant's source-monitoring performance.

This interpretation has also been applied to the differences in performance between recognition tests and source-monitoring tests as a function of response signal deadline. Hintzman and Caulton (1997) found that recognition tests exceed chance performance significantly earlier than source-monitoring tests. In addition, they showed that on recognition tests there is no difference in revival rate as a function of external study sources. Their findings suggest that recognition tests require a smaller amount of information or evidence to determine if an item has been previously studied as compared with the amount of information necessary to make source judgments. Furthermore, this demonstrates that different sources probably share a variety of characteristics, such as semantic, perceptual, and affective details. These shared characteristics may have a sufficient amount of information necessary to make recognition judgments. However, these same characteristics are insufficient or not diagnostic enough to distinguish between two contexts on a source-monitoring test. Therefore, additional time is required to revive information from memory to distinguish between two sources.

The current study is of specific value to the present understanding of the time-course for source-monitoring judgments. Johnson et al. (1994) stated that, "…cognitive operations are either more salient or readily available, or revive more quickly, than perceptual information" (p. 1416). Their findings may be an incomplete picture of the revival rates for source information. Their study suggests that internal sources will be able to revive an adequate amount or type of information for a source judgment at a faster rate compared to perceptual sources. However, the current results suggest the conclusions drawn by Johnson et al. (1994) will have to be modified. Because the revival rate, or slope, changes when one source is tested in different contexts, a source's revival rate does not appear to be a fixed constant but instead appears dependent on the characteristics of the sources that it is paired with at test. More specifically, the amount of time necessary to make a source judgment, should vary as a function of the amount of information required for an individual to discriminate between the two sources. The current findings suggest that the revival rates of internal sources will, in some conditions, be comparable to the revival rates of external sources. For instance, when an internal source is paired with another internal source, the revival rate for the internal sources will be considerably slowed. However, for external-internal source combinations the revival rate for external sources may be better at later deadlines.

In summary, the current set of experiments further identify some of the processes that influence source-monitoring judgments. Some possible future avenues of research include investigating the effects of further varying the degree of similarity between sources and manipulating the number of tested sources. In both cases, it would be predicted that similarity and number of sources should directly affect both the source attribution process and the time necessary to accrue diagnostic information for a given source.

REFERENCES

Batchhelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*, 548-564.

Brown, A. S., Jones, E. M., & Davis, T. L. (1995). Age differences in conversational source monitoring. *Psychology and Aging, 10*, 111-122.

Feller, W. (1961). *Introduction to Probability Theory and Its Applications* (2nd Edition). New York: John Wiley & Sons, Inc.,

Hicks, J. L., & Hancock, T. W. (in Press). Backward associative strength determines source attributions given to false memories. *Psychological Bulletin & Review*.

Hicks, J. L., & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1-15.

Hintzman, D. L., & Caulton, D. A. (1997). Recognition memory and modality judgments: A comparison of retrieval dynamics. *Journal of Memory and Language*, *37*, 1-23.

Hintzman, D. L., Caulton, D. A., & Curran, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 275-289.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences from memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139-154.

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

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1409-1419.

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

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., Foley, M. A., & Kim, J. K. (1982). Pictures and images: Spatial and temporal information compared. *Bulletin of the Psychonomic Society, 19*, 23-26.

Kinchla, R. A. (1994). Comments on Batchhelder and Riefer's multinomial model for source monitoring. *Psychological Review, 101*, 166-171.

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

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

Marsh, R. L., & Bower, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 673-688.

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1137-1151.

Marsh, R. L., & Landau, J. D. (1995). Item availability in cryptomnesia: assessing its role in two paradigms of unconscious plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1568-1582.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 886-897.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 563-582.

Murnane, K., & Bayen, U. J. (1996). An evaluation of empirical measures of source identification. *Memory & Cognition, 24,* 417-428.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803-814.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Tufte, E. (1983) The visual display of quantitative information. Graphics Press; Cheshire, CT.

Reed, A. V. (1973). Speed-accuracy tradeoff in recognition in immediate memory. *Memory & Cognition*, *4*, 16-30.