Journal of Experimental Psychology:
Learning, Memory, and Cognition
1994, Vol. 20, No. 6, 1409-1419

0278-7393/94/$3.00

Time-Course Studies of Reality Monitoring and Recognition

Marcia K. Johnson, John Kounios, and John A. Reeder

Two studies used a response-signal procedure to explore the time course of source-monitoring judgments about perceived and imagined events. Ss judged whether probe words corresponded to pictures that had previously been seen or imagined or were new. Old-new recognition accuracy grew to significant levels before reality-monitoring accuracy, supporting the notion that source monitoring requires more of or a different type of information than does old-new recognition. Also, source identification accuracy developed more quickly for imagined items than for perceived items. This difference in time-course functions is consistent with the idea that memories for perceived and imagined events differ in the relative amounts of various types of information they include (Johnson & Raye, 1981) and that these different types of information may revive or become available to source attribution mechanisms at different rates or may be differentially salient during reality monitoring.

The term *reality monitoring* refers to processes involved in distinguishing memories of external, perceptually based events from memories of internal, reflectively based events (Johnson & Raye, 1981). Reality monitoring is a special case of the broader domain of *source monitoring*—discriminating among memories of various origins (Johnson, 1988; Johnson, Hashtroudi, & Lindsay, 1993). According to Johnson and colleagues (e.g., Hashtroudi, Johnson, & Chrosniak, 1990; Johnson, 1988, 1991a; Johnson & Foley, 1984; Johnson et al., 1993; Johnson & Raye, 1981; Lindsay & Johnson, 1987), source monitoring typically is based on various qualitative characteristics of memories, namely records of perceptual details, contextual and affective information, semantic content, and cognitive operations. *Cognitive operations* refer to the records established at encoding of the cognitive activity involved in perception and reflection (e.g., Johnson, 1983, 1991b; Johnson & Hirst, 1993). Because these various types of information are differentially distributed among memories with different origins, they can be used for source discriminations. These qualitative characteristics are evaluated by decision mechanisms that determine the weighting given to various types of information and the criteria used. For example, decision mechanisms determine whether the amount or type of perceptual detail in a memory for a picture implies that the picture was perceived or that it was only imagined.

This characterization of reality-monitoring processes suggests that people do not remember the source of a memory per se; rather, they infer the source of a memory (Johnson & Raye, 1981; cf. Jacoby, Kelley, & Dywan, 1989). That is, people typically do not directly apprehend the source of a memory by virtue of its being revived. Instead, they make inferences or attributions based on available cues in the form of various types of memorial information. This difference is both epistemologically significant and underlies our working understanding of source monitoring in general (Johnson, 1988; Johnson et al., 1993). It also aids in the generation of testable predictions. For instance, if reality-monitoring processes are based on heuristics rather than on direct detection, they should be systematically error prone; manipulations of the various informational inputs to the heuristic processes should lead to different and predictable patterns of reality-monitoring performance. Consistent with this hypothesis, prior work has shown the general importance of some of these memory attributes for reality monitoring (e.g., Durso & Johnson, 1980; Finke, Johnson, & Shyi, 1988; Intraub & Hoffman, 1992; Johnson, Foley, & Leach, 1988; Lindsay, Johnson, & Kwon, 1991; Rabinowitz, 1989). Other findings support the idea that reality monitoring is based on decision processes that evaluate these attributes according to flexible criteria (e.g., Dodson & Johnson, 1993; Lindsay & Johnson, 1989; Raye, Johnson, & Taylor, 1980).

A further prediction derives from the idea that it may take time for the information that makes up a memory to revive (Johnson et al., 1993). If various memory characteristics revive at different rates and if these characteristics are differentially salient or diagnostic for perceived and imagined source attributions, then subjects may be able to identify the origin of one type of event at earlier stages of processing than at another. In our studies, we explore this possibility by investigating the time course of reality-monitoring judgments by using a response-signal methodology (Dosher & Rosedale, 1991; Pachella, 1974; Wickelgren, 1977; see also Meyer, Irwin, Osman, & Kounios, 1988). With this procedure, subjects are signaled to respond at varying delays after the presentation of a test item. Accuracy can then be examined as a function of the time available for processing.

In addition to allowing us to compare time-course functions

Marcia K. Johnson and John A. Reeder, Department of Psychology, Princeton University; John Kounios, Department of Psychology, Tufts University.

We thank Evan Chua-Yap for conducting relevant pilot studies, Bill Prinzmetal for advice on the programming in Experiment 1, and Steve Lindsay for comments on an early write-up of Experiment 1. This research was supported by National Institute on Aging Grant AG09253 to Marcia K. Johnson and National Institute of Mental Health Grant CEP 1 R29 MH45447-01A3 to John Kounios.

Correspondence concerning this article should be addressed to Marcia K. Johnson, Department of Psychology, Princeton University, Princeton, New Jersey 08544. Electronic mail may be sent to mkj@clarity.princeton.edu,

for identifying the origin of perceived and imagined events, the present design and analysis also permit a comparison of the time-course functions for reality monitoring and old–new recognition. There are several types of evidence consistent with the general idea that old–new recognition and source monitoring are sometimes based on different kinds of information or processes (e.g., Johnson & Raye, 1981). For example, manipulations do not always affect the two measures in the same way, and different subject groups may be equated on recognition but show differences in source monitoring (Ferguson, Hashtroudi, & Johnson, 1992; Johnson & Raye, 1981; Johnson et al., 1993; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991). Such dissociations arise because old–new recognition can take place on the basis of almost any kind of information, whereas specific types of information are necessary for source monitoring (cf. Raye, 1976). All that may be needed for recognition is some differential response to old and new items. Thus old items may have an advantage over new in frequency information (Underwood, 1972), perceptual fluency (Jacoby & Dallas, 1981), echo intensity (Hintzman, 1988), amount of activation of individual representations (Mandler, 1991), or amount of associative spread of activation (Gillund & Shiffrin, 1984). In contrast, source monitoring typically requires more differentiated information (Johnson et al., 1993). Differentiation can be distinguished from ideas such as strength or activation level. Differentiated information has phenomenal qualities such as color, form, spatial position, and so forth. That is, for source monitoring it matters not only how much information is available but also what the information is.

In short, decision mechanisms can use relatively undifferentiated information for old–new recognition decisions but require more differentiated information for source-monitoring decisions. We have further hypothesized that it may take time for the information that makes up a memory to become differentiated (Johnson et al., 1993). Consistent with this idea is the finding that response times to make old–new decisions are typically faster than the response times to make source-monitoring decisions (Johnson, Foley, & Chua-Yap, 1983, described in Johnson, 1985). Thus, sufficient information appears to be available for old–new discrimination before sufficient information becomes available for source monitoring. However, in Johnson et al. recognition and source-monitoring response times were obtained for separate groups of subjects. The response times for old–new subjects might have been faster than those of source-monitoring subjects only because subjects making old–new decisions were making responses involving two alternatives (old and new), whereas subjects making source-monitoring judgments were making responses involving three alternatives (perceived, imagined, and new). The present experiments allowed us to compare old–new recognition and source monitoring with number of response options held constant.

In our experiments, accuracy measures for both old–new and source discriminations were obtained from the same subjects, for the same items, at various probe-signal lags. The major task from the subjects' point of view was source monitoring, and the measure of old–new discrimination was derived by assessing subjects' recognition accuracy independently of their source-monitoring accuracy by using a multinomial-modeling approach recently described by Batchelder and Riefer (1990), which is summarized below (see also Appendix). Thus, the present design permits a stronger test of the hypothesis that the overall time course of source monitoring might be slower than the time course for old–new discriminations than would be afforded by a simple comparison of response times.

To summarize, the primary questions of interest here were as follows: (a) Do reality-monitoring judgments for perceived and imagined events show similar or different time-course functions? and (b) Do reality monitoring and recognition have similar or different time-course functions?

Experiment 1

Processing time was manipulated by varying the temporal lag between presentation of the test stimulus and the onset of an auditory tone that signaled subjects to make their responses immediately. Response–signal lag values of 300, 400, 500, and 1,500 ms were chosen to allow an examination of a broad temporal region of the time courses of reality monitoring and recognition. In an additional response time condition, processing time was controlled by the subjects themselves, who were instructed to respond to each test stimulus as quickly as possible while maintaining high accuracy. This condition should yield performance levels typical of previous source-monitoring experiments.

Method

Subjects. Experiment 1 included 72 subjects in the response–signal conditions (n = 32 received response–signal lags of 300 and 400 ms randomly intermixed, and n = 40 received lags of 500 and 1,500 ms randomly intermixed, as described below), and 24 subjects were in the response time condition.[1] All subjects were Princeton and Tufts undergraduates who were tested individually and received either nominal payment or partial course credit for their participation.

Apparatus. Acquisition items were presented with a slide projector. The test phases of the experiments were controlled by an Apple IIe microcomputer equipped with a Mountain (millisecond) computer clock. Subjects viewed the test stimuli on the computer's video monitor and responded by pressing buttons on the keyboard.

Stimuli. The stimuli were selected from the standardized set of black-and-white line drawings of common objects published by Snodgrass and Vanderwart (1980). One hundred of these pictures were selected on the basis of two criteria. First, each picture had to have a one-word label to describe it. Second, these one-word labels had to be specific and concrete enough so that they could evoke a corresponding mental image. This second criterion was applied according to the intuitions of the experimenters.

Procedure. During the acquisition phase, subjects were presented with 50 single-word object labels. After half of these, subjects saw a picture corresponding to the label. The other half of the trials consisted of the presentation of a verbal label followed by a blank screen signaling the subject to imagine a picture corresponding to the label. During the test phase, all of these previously viewed word labels

[1] The subjects were not assigned to these groups randomly; the three groups corresponded to three separate experiments. However, the data are presented here as a single experiment for reasons of clarity and because the important conclusions were replicated in Experiment 2.

were displayed again, randomly mixed with 50 new word labels. The selection of old and new word labels was counterbalanced across subjects, as was the selection of word labels corresponding to perceived and imagined pictures.

Subjects were told that the first part of the experiment was a test of aesthetic perception and visual imagination. During the acquisition phase, each subject sat in front of a screen upon which slides of the pictures and word labels were presented. On each trial, a word label was displayed for 2 s, followed by either a corresponding picture or a blank illuminated screen (for 6 s). Subjects were told to rate each picture for artistic merit. For trials with blank screens, they were told to form a visual image of a black-and-white line drawing corresponding to the verbal label and to rate that imagined drawing for artistic merit. Subjects recorded their rating judgment by writing a number from one to five on a score sheet.

Following the acquisition phase, subjects were told that they were about to participate in another brief, unrelated experiment. They were then taken to a different room and were seated at a microcomputer. At this point, subjects in the response–signal condition received practice at trading speed for accuracy by participating for approximately 15 min in a standard short-term memory-scanning task (e.g., Sternberg, 1969), incorporating the response–signal technique (see Reed, 1976). The response–signal lags used in this practice task were identical to the lags used in the subsequent test phase. Subjects in the conventional response time condition received practice at the memory-scanning task without the response signals.

Next came the test phase of the experiment. Subjects were again seated at the computer and were told that they would see a sequence of words displayed on the screen in front of them. For each of these words, they had to judge whether the word corresponded to a previously perceived picture or a previously imagined picture or whether it was a new item that had not previously been encountered in the experiment. Each test stimulus was preceded by a 500-ms warning stimulus (a plus sign). Subjects in the response–signal condition were asked to wait until they heard an auditory tone (of 100-ms duration) before responding, but to respond immediately on hearing the tone. Thirty-two subjects were tested at short signal lags of 300 and 400 ms, and 40 were tested at longer lags of 500 and 1,500 ms. For any given subject, lag was randomly varied across trials. Any response made before 75 ms had elapsed since the onset of the response signal elicited a "TOO FAST!" message on the monitor. Any response made more than 300 ms after the onset of the response signal elicited a "TOO SLOW!" message. No error feedback was provided. Subjects were instructed that speed of response was the highest priority and accuracy was the second priority. In the response time condition, subjects were simply asked to respond as quickly and accurately as possible.

All subjects were instructed to use three keys on the computer's keyboard to indicate whether each probe stimulus corresponded to a picture that was previously perceived, imagined, or had not previously been viewed. The one, dash, and spacebar keys were used to indicate these responses. Subjects were instructed to depress the one key with the left index finger, the dash key with the right index finger, and the spacebar with the thumb of the dominant hand. The response mapping of the one and dash keys was counterbalanced across subjects between the perceived and imagined responses, whereas the spacebar was always used to indicate that a probe stimulus was new. Each test stimulus remained on the screen until a legal response was made. The intertrial interval was 2,500 ms.

To reduce the contribution to accuracy measures of response–signal trials on which subjects may have "cheated" by taking more than the allowed processing time, the analyses reported below were performed only on data from trials on which the subject responded before 500 ms had elapsed from the onset of the response signal. This filtering resulted in about 16% of the responses from the response–signal condition being discarded; for the 300, 400, 500, and 1,500 ms lags, 30%, 25%, 25%, and 9% of the responses were discarded, respectively, equally distributed across perceived and imagined items. None of the responses from the response time condition were discarded.

Results and Discussion

The data were sorted into 3 × 3 tables containing the frequencies with which each of the source-monitoring responses (perceived, imagined, and new) were given to items from each of the three sources (P, I, and N). Table 1 reports these frequencies separately for the four lags in the response–signal condition and the response time condition.

The frequency tables were subsequently analyzed for old–new recognition and source discrimination performance according to the multinomial-processing models described by Batchelder and Riefer (1990; see also Appendix). These are high-threshold simplifications of signal detection theory (Green & Swets, 1966). We used their Model 5b to estimate several parameters. The goodness of fit of Model 5b for these data and those reported for Experiment 2 was assessed with loglikelihood ratio tests; the test statistic G^2 is distributed like χ^2 (1). With two exceptions, G^2 values testing the fit of Model 5b obtained from each 3 × 3 frequency matrix ranged from 0.23 to 2.19, indicating a reasonable fit for Model 5b. The exceptions were a $G^2 = 4.96$ for the 300-ms processing interval and a $G^2 = 40.49$ for the 500-ms processing interval in Experiment 1. However, collapsed across lag, Model 5b was a good fit for Experiment 1, $G^2 = 1.80$. In addition, the general pattern of results from Experiment 1 was replicated in Experiment 2 and thus, for clarity of exposition and consistency, we used Model 5b throughout.

The parameter D estimates old–new detection, the probability of successfully discriminating old from new items. D was collapsed across Source P and Source I stimuli because old–new recognition did not differ for P and I items in either

Table 1
Frequencies of Perceived, Imagined, and New Responses for Each of the Three Item Sources in the Response–Signal and RT Conditions of Experiment 1

	Signal lag														
	300 ms			400 ms			500 ms			1,500 ms			RT		
Source	P	I	N	P	I	N	P	I	N	P	I	N	P	I	N
Perceived	76	71	132	106	78	109	169	102	102	368	47	45	480	89	31
Imagined	69	69	137	77	110	115	114	169	87	81	331	40	29	545	24
New	85	47	555	62	41	625	268	48	573	180	79	689	5	19	1,174

Note. P = perceived response; I = imagined response; N = new response; RT = response time.

Table 2
Multinomial Model Parameter Estimates for Experiment 1 (With 95% Confidence Intervals)

Parameter	300 ms		500 ms		900 ms		1,500 ms		RT		Lag conditions combined	
	M	*CI*	*M*	*CI*	*M*	*CI*	*M*	*CI*	*M*	*CI*	*M*	*CI*
D	.40	(.34–.46)	.56	(.51–.61)	.60	(.55–.66)	.87	(.84–.90)	.95	(.94–.96)	.64	(.94–.96)
d_1	.00	(−.31–.31)	.00	(−.31–.31)	.00	(−.25–.25)	.65	(.53–.78)	.80	(.75–.86)	.00	(−.16–.16)
d_2	.17	(−.06–.41)	.33	(.15–.51)	.56	(.46–.67)	.75	(.68–.81)	.76	(.55–.97)	.61	(.56–.66)
b	.19	(.12–.27)	.14	(.05–.23)	.36	(.32–.40)	.27	(.22–.33)	.02	(−.14–.18)	.25	(.22–.28)
g	.58	(.55–.61)	.59	(.56–.61)	.74	(.71–.78)	.70	(.67–.72)	.21	(.20–.22)	.72	(.70–.73)

Note. D = estimates of old–new recognition; d_1 = source identification for perceived items; d_2 = source identification for imagined items; b = bias for responding "old" to a nondetected item; g = bias to label items as originating from Source P, combined across nondetected stimuli and detected but nondiscriminated stimuli; RT = response time; lag = response–signal lag; *CI* = confidence interval.

experiment (see Appendix). Items that are not successfully detected are guessed to be old with probability b. The parameters d_1 and d_2 give the probability of successful source identification for P and I items, respectively. The bias parameter g gives the probability that successfully detected old items that were not discriminated according to source, as well as any items that were not detected but which were guessed to be old, are labeled as originating from Source P. Note that for the three sensitivity parameters D, d_1, and d_2, chance performance is indicated by a value of zero. Values of b above .5 indicate a bias to guess old and values below .5 a bias to say new. Values of g above .5 indicate a bias to guess perceived and values below .5 a bias to say imagined. Additional discussion of this technique and its application to our data is provided in the Appendix.[2]

Table 2 gives the parameter estimates yielded by Model 5b for the conditions in Experiment 1, along with their 95% confidence intervals,[3] and Figure 1 shows the corresponding time-course functions. (There are several precedents for obtaining time-course functions across groups in which any particular subject encounters only a subset of the processing intervals, for example, Pachella, 1974.) Table 2 also includes parameter estimates and confidence intervals for the data collapsed across the four signal lags in the response–signal conditions to make overall comparisons easier. The main findings are clear from Figure 1 in combination with the confidence interval information given in Table 2.

First, as is evident from Figure 1, both old–new recognition (D) and source monitoring (d_1 and d_2) improved as more processing time was available. (Compare, for example, accuracy at 300 ms with accuracy at 1,500 ms for D, d_1 and d_2 in Table 2.) Old–new recognition accuracy was above the chance value of zero even at the shortest lag, 300 ms, and in general exceeded source-monitoring accuracy (see last column in Table 2). With respect to reality monitoring, overall, subjects were better able to identify the origin of I items than P items. In addition, it is striking that subjects' source identification of I items was significantly above chance by 400 ms but did not exceed chance until the 1,500-ms condition for the P items. It is also notable that even though information that helps identify items as perceived was initially less available than information subjects used to identify items as imagined, it evidently caught up in the later regions of the time-course functions.

Low (below .5) values for the bias parameter b in Table 2 indicate bias to say "new" rather than "old." Estimates of g above .5 indicate some tendency for subjects to say "P" rather than "I" in all but the response time group in which the bias was to say "I" rather than "P."

The difference in time-course functions for old–new recognition and reality monitoring is consistent with the notion that these two discrimination judgments rely on different amounts or types of memorial information. In addition, the overall pattern shown in Figure 1 constitutes preliminary evidence of different time courses not only for recognition and reality-monitoring judgments but for source discrimination judgments about perceived and imagined items. If further substantiated, these findings would support the notion that rather than

[2] It is important to note that the model depends on assumptions that are mathematically necessary for obtaining the several parameters, but which may only approximate the psychological organization of the processes underlying reality monitoring. For example, the processing-tree model (Batchelder & Riefer, 1990) includes a sequence of discrete states: Old–new detection is followed by source discrimination and then by bias effects. In contrast, the memorial information informing recognition may be revived in parallel to source-relevant information. Moreover, decision mechanisms may not be in one discrete state or another. The details associated with a remembered event may be more or less representative of a particular source and produce a continuum of judgment certainty (e.g., Johnson & Raye, 1981; Johnson et al., 1993). Nevertheless, the fact that recognition accuracy developed more quickly than reality-monitoring accuracy (see *Results and Discussion* section) is consistent with Batchelder and Riefer's multinomial model, which puts old–new detection at the top of the processing tree.

[3] These confidence intervals are obtained from asymptotic approximations, which were described in general by Riefer and Batchelder (1988). As with confidence intervals yielded by normal approximations to the binomial, it is possible for their values to stray below zero or above unity. Furthermore, it should be noted that because some parameters in multinomial models of this kind modulate the effects of others, it is not unusual to observe widths for those intervals that vary markedly from parameter to parameter (we thank W. Batchelder for pointing this out).

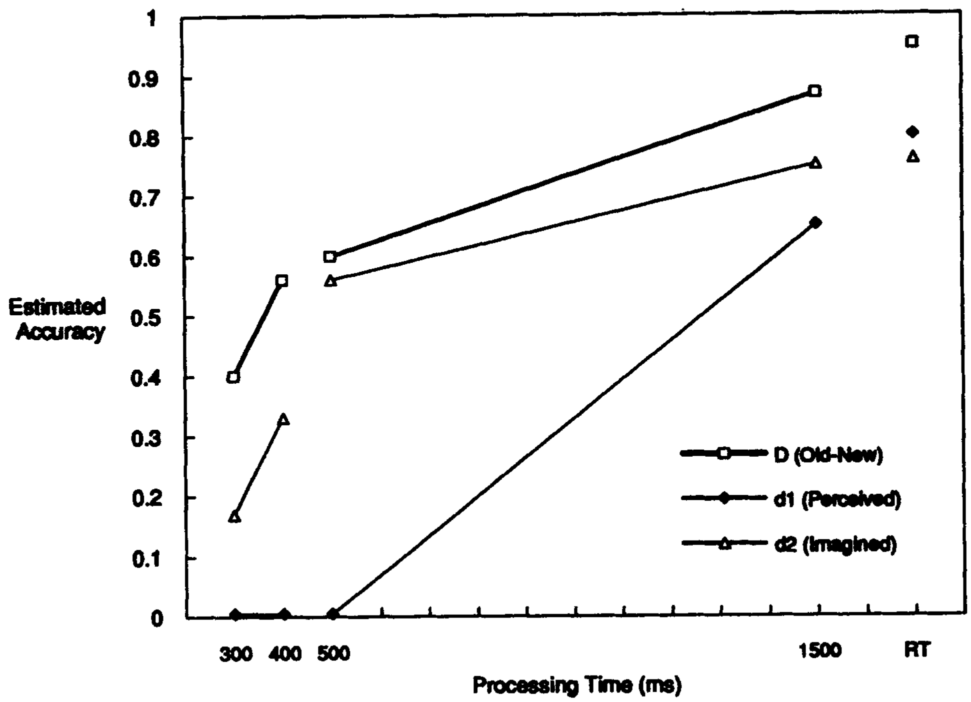

Figure 1. Composite time-course functions constructed by combining accuracy estimates for different response–signal lags from Experiment 1. RT = reaction time.

being based on the accumulation of a single type of information evaluated against some singular criterion, reality monitoring depends on the output of a decision mechanism or set of separate mechanisms that operate on different types of information.

Experiment 2

There are two general procedures for obtaining pictures of the time courses of psychological processes. In Experiment 1 we used the method of measuring different subjects' performance after different amounts of processing time and combining these measures into functions that specify accuracy at each of the various intervals. Experiment 2 presented the alternate technique of gathering accuracy measures across the same range of processing times for all subjects. It is worth demonstrating that similar outcomes were yielded by these two methods because there are advantages for each. For example, with limited numbers of stimulus items or limited time to test any particular subject, the first method could have the advantage of more measures per interval. On the other hand, the second, within-subject method allows greater confidence that attention and motivation are equated across processing intervals.

Method

At acquisition 80 single-word picture labels were presented in random order. Forty of these were P (perceived) items; they were each followed by the presentation of a corresponding picture enclosed in a circular frame. Another 40 items were I (imagined) items; they were each followed by an empty circular frame in which subjects were instructed to imagine a picture of the named object. At test, these 80 items were mixed with 40 N (new) items and presented in random order as probes. Speeded responses (P, I, or N) to these probes were collected at four response–signal lags. Thus, there were 40 test trials for each of the three item types (perceived, imagined, and new), and these were subdivided equally among the four lag conditions (300, 500, 900, and 1,500 ms), yielding 10 trials per condition per subject. Items were counterbalanced across the 12 conditions.

Subjects. Twenty-five paid undergraduate volunteers at Princeton University participated in Experiment 2. One subject was dropped from the experiment for failing to follow the instructions appropriately. Each subject was tested individually.

Apparatus. The acquisition, practice, and test phases were all conducted using an IBM-compatible microcomputer. Stimuli were presented on a 16″ color monitor, and responses were made by using a standard keyboard.

Stimuli. One hundred and twenty pictures with single-word labels were selected from the Snodgrass and Vanderwart (1980) set for use in the acquisition and test phases. For the response–signal practice phase, described below, 40 two-syllable words and 20 paralogs (two-syllable nonwords) were selected from the list provided by Taylor and Kimble (1967). Although selection of these items was essentially random, words were accepted only if, in the judgment of the experimenters, they were unrelated to the picture labels used as probe stimuli in the acquisition and test phases. Additionally, paralogs considered to be easily mistakable for words were rejected. All materials were displayed on the computer screen in dark blue against a pale green background. Text was approximately ¾ in. high.

Procedure. The acquisition, practice, and test phases were each preceded by a set of instructions displayed by the computer. The instructions for the acquisition phase informed subjects that they would be performing a task requiring them to look at and to imagine drawings of objects. Each of the 80 trials consisted of a 1.5-s presentation of a single-word picture label, followed by a 6-s display containing a circle of 4.25 in. diameter. On the P trials, the circle enclosed the picture corresponding to the label; on the I trials, the circle was empty. Subjects were instructed to look at the picture in the first case and to imagine a picture of the named object in the second case. Imagined pictures, they were told, should be line drawings like the ones they saw during the perceived-item trials. Furthermore, subjects were encouraged to envision the imagined pictures being enclosed in the circle, mentally projecting them onto the screen.

While each circular frame (with or without a picture) remained on-screen, subjects were engaged in an orienting task that was designed to encourage both perceptual encoding of the displayed objects and intentional imagination of the nondisplayed objects: They were asked to think about how well the picture (or image) illustrated the object in question. Artistic merit and clarity of representation were suggested as criteria for the judgment, and subjects were instructed to attempt to apply these equivalently to the perceived and imagined pictures. When the circle disappeared after 6 s, the instruction "Rate the drawing" was displayed, along with the following options: (1) Good, (2) Adequate, or (3) Poor, signaling the subject to press the key corresponding to their decision about the merit of the perceived or imagined illustration. No time limit for responding was enforced, but subjects were encouraged to respond quickly to keep the task moving along.

Following completion of the acquisition phase, subjects were informed that they would be performing a very different sort of task but were not told that it was only meant to be practice for the test phase. The practice phase was actually a special lexical decision task that collected responses for the four response–signal lags (300, 500, 900, and 1,500 ms) to be used in the upcoming test phase (described below). In 60 trials subjects identified verbal items as uppercase words, lowercase words, or nonwords. There were 20 items of each type, assigned equally to the four lags. This triplet of item types and the three-alternative decision were designed to be analogous to the conditions of the P–I–N judgment involved in the upcoming reality monitoring and recognition test. The entire practice phase lasted approximately 10 min.

In the test phase subjects were asked to make reality-monitoring judgments. There were 120 trials, each consisting of a 500-ms fixation stimulus (a plus sign), followed by presentation of the probe word, and then by the auditory-response signal. Stimulus order, and therefore lag, was randomly determined. Subjects were instructed to indicate whether each displayed word corresponded to a previously perceived picture, a previously imagined picture, or was a new word for the experiment. Responses were made using the keyboard: A finger from each hand made the P and I responses by using the W and left bracket keys (counterbalanced across subjects), whereas the dominant thumb was used to make N responses on the spacebar.

Responses were signaled by a 100-ms tone generated by the computer after 300, 500, 900, or 1,500 ms had elapsed since stimulus onset. Subjects were encouraged to be as accurate as possible but were instructed to make their responses immediately on hearing the tone. As in Experiment 1, responses made less than 75 ms after the onset of the response signal resulted in a "TOO FAST" message, responses made more than 300 ms after signal onset resulted in a "TOO SLOW" message, and there was no error feedback. Each item was displayed until a legal response was made, or until 1,000 ms had elapsed since response-signal onset. The intertrial interval was 2,500 ms. Altogether, the entire experiment could be completed in about 45 min.

Results and Discussion

Data collected during the test phase were filtered according to the method applied in Experiment 1, which disqualified responses made before response–signal onset or more than 500 ms after response–signal onset. Overall, 18% of the responses were discarded in this way—31%, 13%, 11%, and 18% for 300, 500, 900, and 1,500 ms lags, respectively (responses to P and I items were not differentially lost).

First, the data were sorted into 3 × 3 tables containing the frequencies with which each response was made to each item type, for the four signal lag values. These frequencies are reported in Table 3. Next, the frequency tables were subjected to the multinomial analysis described earlier. (As noted previously, Model 5b provided a good fit at all response lags in Experiment 2.) The resulting parameter estimates (with their 95% confidence intervals) are given in Table 4. The corresponding time-course functions are shown in Figure 2.

As can be seen in Figure 2 and verified by examining the confidence intervals in Table 4, old–new recognition accuracy was generally greater than source-monitoring accuracy, and subjects were generally more accurate identifying the source of imagined than perceived items. Furthermore, the source-monitoring accuracy exceeded chance for imagined items at 300 ms but not for perceived items until 500 ms. The low (below .5) bias parameter b in Table 4 indicates that subjects were biased to say "new" rather than "old." Estimates of g above .5 indicate a slight bias (in the early lags only) to say "P"; the bias was to say "I" at 1,500 ms.

The most important finding, illustrated by a comparison of Figures 1 and 2, is that the general pattern obtained in Experiment 1 was replicated in Experiment 2. Old–new recognition and reality-monitoring discrimination of both perceived and imagined items all exhibited visible upward trends with longer processing times, and recognition was superior to reality monitoring. Moreover, again subjects could identify items as previously imagined before they demonstrated any

Table 3
Frequencies of Perceived, Imagined, and New Responses for Each of the Three Item Sources in Experiment 2

	Signal lag											
	300 ms			500 ms			900 ms			1,500 ms		
Source	P	I	N	P	I	N	P	I	N	P	I	N
Perceived	65	48	35	135	43	23	169	31	12	176	22	5
Imagined	32	79	48	31	139	31	24	184	6	10	175	10
New	23	14	150	7	6	210	3	3	208	1	2	187

Note. P = perceived response; I = imagined response; N = new response.

Table 4
Multinomial Model Parameter Estimates for Experiment 2 (With 95% Confidence Intervals)

	300 ms		500 ms		900 ms		1,500 mg		Lag conditions combined	
Parameter	*M*	*CI*	*M*	*CI*	*M*	*CI*	*M*	*CI*	*M*	*CI*
D	.66	(.60–.73)	.86	(.82–.89)	.96	(.94–.98)	.96	(.94–.98)	.88	(.86–.90)
d_1	.00	(−.45–.45)	.48	(.14–.82)	.69	(.42–.96)	.83	(.69–.98)	.51	(.35–.68)
d_2	.56	(.35–.77)	.67	(.46–.87)	.77	(.57–.97)	.84	(.56–1.11)	.76	(.69–.83)
b	.20	(.05–.35)	.06	(−.21–.33)	.03	(−.37–.43)	.02	(−.52–.55)	.07	(−.05–.20)
g	.59	(.53–.64)	.54	(.51–.57)	.50	(.48–.52)	.33	(.32–.35)	.58	(.56–.59)

Note. D = estimates of old–new recognition; d_1 = source identification for perceived items; d_2 = source identification for imagined items; b = bias for responding "old" to a nondetected item; g = bias to label items as originating from Source P, combined across nondetected stimuli and detected but nondiscriminated stimuli; lag = response–signal lag; *CI* = confidence interval.

sensitivity to the source of previously perceived items. Although the confidence intervals around particular estimates of d_1 and d_2 tended to be larger in Experiment 2, the replication of the overall pattern of differences in discriminability between external and internal sources supports its reliability.

Subjects in typical response–signal studies receive many more trials than used here, often over multiple sessions on a task, and hence are highly practiced at responding to the signal. We limited the number of trials within a session because imagination is effortful and we wanted subjects to maintain a high degree of involvement over the acquisition trials; we did not test subjects for multiple sessions because we were interested here in performance characteristics when subjects did not know the nature of the upcoming test during the acquisition phase. Our results are especially encouraging because they indicate that systematic data can be obtained with relatively few observations by using both completely within-subject designs and designs in which each subject only serves in a subset of the processing intervals.

It should also be noted that overall accuracy levels were higher throughout the time course in Experiment 2 compared with Experiment 1. For example, discrimination of imagined items at 300 ms was .17 in Experiment 1, compared with .56 in Experiment 2; old–new recognition accuracy rose to .87 at 1,500 ms in Experiment 1, but to .96 in Experiment 2. There are a few possible sources of this general superiority. For example, the orienting task differed somewhat between Experiments 1 and 2. Also, the practice session in Experiment 2 was substantially shorter than that in Experiment 1, leaving a shorter retention interval between acquisition and test. Fi-

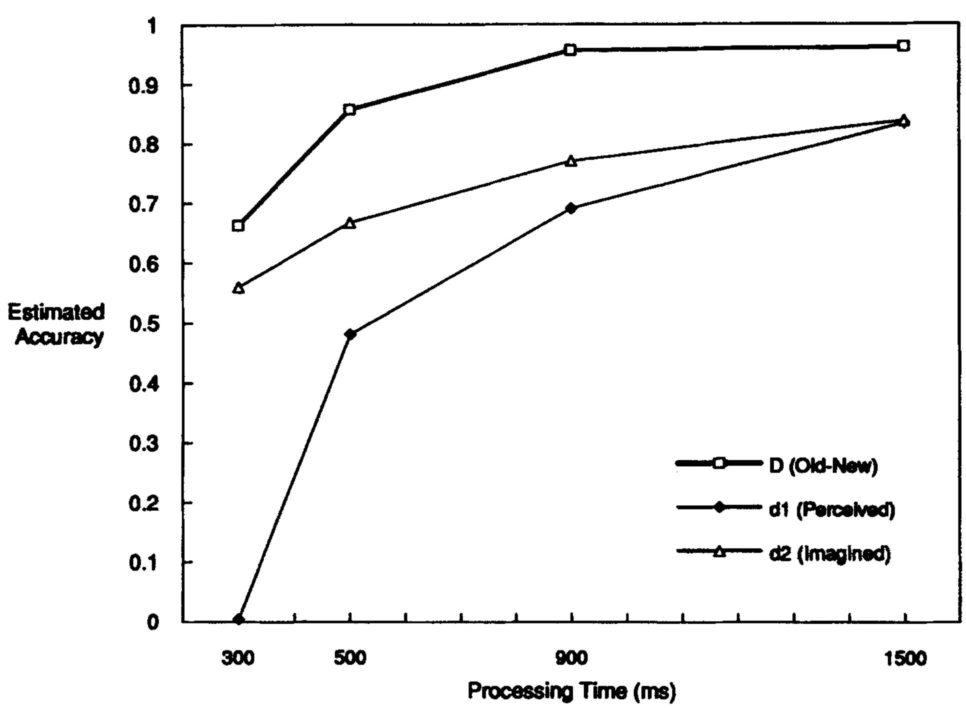

Figure 2. The time-course functions for old–new recognition and reality-monitoring performance in Experiment 2.

nally, the equipment used was different in the two experiments and the experimental conditions of the acquisition and test phases were more similar in Experiment 2; they were conducted in the same room with all stimuli presented on the computer screen in all phases. In any event, it is noteworthy that the pattern of discrimination scores was similar at both lower and higher levels of performance. For example, even with the higher overall accuracy estimates in Experiment 2, subjects still showed no sensitivity to the source of previously perceived items at 300 ms, when they could already identify previously imagined items with substantial accuracy.

General Discussion

In our experiments, subjects saw some pictures and imagined others and were later presented with word probes and asked to specify whether each item had been previously seen, imagined, or was new. The lag between the probe word and a signal to respond was systematically varied, including values of 300, 400, 500, 900, and 1,500 ms and an unconstrained response time condition. The major findings from these experiments are evident in Figures 1 and 2. First, as the response–signal lag increased, both old–new recognition and source identification improved. Significant levels of information indicating past occurrence, however, were clearly available before information specifying source. At a signal lag of 300 ms, subjects in Experiment 1 exhibited greater discrimination between old and new items than between perceived and imagined items. This finding is consistent with the idea that recognition of an event's prior occurrence can be based on different information than that specifying source or can be based on a less differentiated form of the same information (Johnson & Raye, 1981; see also Johnson et al., 1993; Raye, 1976). Of course, as increasingly differentiated information becomes available, we would expect some of the same information that permits source discriminations to be used in old–new recognition judgments. For example, subjects might be especially confident that they had seen an item previously if they experienced not just a feeling of recent familiarity with the concept but if they also remembered a specific line drawing of the item. Thus the fact that old–new recognition and source monitoring can be dissociated at early lags as they were here is not an argument that they should necessarily always be dissociated (cf. Johnson et al., 1993).

A second major finding was that varying the lag between the probe word and the response signal produced different time-course functions for reality-monitoring judgments about previously perceived and imagined pictures. Subjects could identify imagined items as having been imagined well before they could identify perceived items as previously perceived. In our data, above-chance source identification of imagined items was present at the 400-ms signal lag in Experiment 1 and at the 300-ms signal lag in Experiment 2, but above-chance source identification for perceived items was not present until 1,500 ms in Experiment 1 and 500 ms in Experiment 2. We are not placing any particular emphasis on the absolute values of the parameters, which would be expected to vary with a number of factors (e.g., acquisition task, practice in the response–signal task, retention interval, and so forth); rather, the overall pattern obtained is of interest. The clearly different shapes of the time-course functions for perceived and imagined events are consistent with the idea that the memory representations of these events include different types of information or different distributions of the same types of information (Johnson et al., 1993; Johnson & Raye, 1981).

The source-monitoring framework proposed by Johnson et al. (1993; see also Johnson & Raye, 1981) provides a context for interpreting the different shapes of these time-course functions. According to this framework, memories for imagined and perceived events typically differ in the distributions of the various types of characteristics they include (i.e., perceptual, contextual, semantic, affective, and cognitive-operations information). In our experiments (and with such simple materials), we would not expect much difference between perceived and imagined events in contextual, semantic, and affective information. However, two differences might be especially important: Memories for imagined events would be expected to include more available information about the cognitive operations that established them, and memories for perceived events would be expected to include more perceptual detail. If so, the differences in the shapes of the time-course functions for source judgments about items from these two sources would suggest that information about cognitive operations is either more salient or readily available, or revives more quickly, than perceptual information (at least given the types of encoding conditions and test probes used here).

The fact that the time-course functions for old–new recognition of perceived and imagined items did not differ significantly suggests that such discriminations were based on information that becomes available at about the same rate for imagined and perceived events. We confirmed this finding in an independent study that included response–signal lags of 300 and 400 ms and required subjects simply to respond "old" or "new" to each item. Subjects' ability to differentiate P from N items did not differ significantly from their ability to differentiate I from N items, as indexed by d', nor were the estimates of beta different for P and I items. Unlike the memorial evidence informing reality monitoring, then, recognition appears to have relied on the revival of information that is similarly represented in memories for perceived and imagined events. In this situation, this equivalence would likely consist of some combination of information specifying the prior activation of underlying concepts relevant to the item (i.e., semantic information) and increased fluency of perceptual processing (e.g., Jacoby & Dallas, 1981) of the visually presented probes (because acquisition items were designated with visually presented words on both perception and imagination trials). Again, the presence of significant old–new discrimination at very short response–signal lags suggests that such information is available very quickly or easily and yields a feeling of familiarity.

Overall, these experiments demonstrated the usefulness of the time-course approach for exploring memory characteristics and decision mechanisms involved in monitoring the source of information in memory. Some planned future studies are aimed at clarifying the contribution to old–new recognition and reality monitoring made by the revival of information with different qualitative characteristics. For example, changing presentation modalities of the concept cues between acquisition and test, varying the amount of cognitive operations required by the acquisition task, and manipulating the degree to which test probes reinstate the perceptual details and cognitive operations recorded during encoding could selectively affect the time-course functions for recognition and reality-monitoring decisions about perceived and imagined events. Furthermore, it should be possible to affect the pattern of results by influencing the criteria people use to evaluate the information that makes up their memories. Additional demonstrations of the independence and manipulability of the time-course functions derived here would provide evidence about the cues used to distinguish between memories from external and internal sources. The response–signal technique could also be used to explore types of source monitoring in addition to reality monitoring. For example, it could be used to compare the relative availability of location, color, speaker, and temporal information about events derived from various external sources (Johnson et al., 1993).

References

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

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

Dodson, C. S., & Johnson, M. K. (1993). The rate of false source attributions depends on how questions are asked. *American Journal of Psychology, 106,* 541–557.

Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language, 30,* 125–160.

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

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 7,* 443–452.

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

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Greeno, J. G., & Steiner, T. E. (1964). Markovian processes with identifiable states: General considerations and application to all-or-none learning. *Psychometrika, 29,* 309–333.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging, 5,* 119–126.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Hu, X. (1990). *Source monitoring program.* [Version 1.0]. Irvine, CA: University of California, Institute for Mathematical Behavioral Sciences, School of Social Sciences.

Intraub, H., & Hoffman, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. *American Journal of Psychology, 105,* 101–114.

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 III & F. I. M. Craik (Eds.), Varieties of memory and consciousness: Essays in honour of Endel Tulving (pp. 391–422). Hillsdale, NJ: Erlbaum.

Johnson, M. K. (1983). A multiple-entry, modular memory system. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 17, pp. 81–123). San Diego, CA: Academic Press.

Johnson, M. K. (1985). The origin of memories. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 4, pp. 1–27). San Diego, CA: Academic Press.

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General, 117,* 390–394.

Johnson, M. K. (1991a). Reality monitoring: Evidence from confabulation in organic brain disease patients. In G. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury* (pp. 176–197). New York: Oxford University Press.

Johnson, M. K. (1991b). Reflection, reality-monitoring, and the self. In R. Kunzendorf (Ed.), *Mental imagery* (pp. 3–16). New York: Plenum.

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues, 40,* 33–50.

Johnson, M. K., Foley, M. A., & Chua-Yap, E. (1983). [Response times for source monitoring and recognition of large and small pictures]. Unpublished raw data.

Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition, 16,* 337–342.

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

Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 241–286). Hillsdale, NJ: Erlbaum.

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

Lindsay, D. S., & Johnson, M. K. (1987). Reality monitoring and suggestibility: Children's ability to discriminate among memories from different sources. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 92–121). New York: Springer-Verlag.

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

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52,* 297–318.

Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 207–225). Hillsdale, NJ: Erlbaum.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review, 95,* 183–237.

Pachella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.

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

Raye, C. L. (1976). Recognition: Frequency or organization? *American Journal of Psychology, 89,* 645–658.

Raye, C. L., Johnson, M. K., & Taylor, T. H. (1980). Is there something special about memory for internally generated information? *Memory & Cognition, 8,* 141–148.

Reed, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory & Cognition, 4,* 16–30.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339.

Riefer, D. M., Hu, X., & Batchelder, W. H. (1994). Response strategies in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 680–693.

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.

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.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Psychologist, 57,* 421–457.

Taylor, J. D., & Kimble, G. A. (1967). The association value of 320 selected words and paralogs. *Journal of Verbal Learning and Verbal Behavior, 6,* 744–752.

Underwood, B. J. (1972). Word recognition memory and frequency information. *Journal of Experimental Psychology, 94,* 276–283.

Wickelgren, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.

Appendix

Multinomial Modeling of Source-Monitoring Data

The analyses of source-monitoring data described here are based on the multinomial-modeling approach recently described by Batchelder and Riefer (1990). This technique uses high-threshold versions of the corresponding signal detection models (Green & Swets, 1966) to estimate separate sensitivity and bias parameters for old–new and source judgments. Batchelder and Riefer's multinomial approach is based on breaking down response selection into components that can be organized into a processing-tree model. As Figure A1 shows, the first step (logically speaking, and not necessarily in the sense of an information-processing stage with specific real-time properties) is old–new detection (i.e., recognition). An item from either Source A or Source B can successfully be judged to be old or new. The probability of successful detection is D_1 for items from Source A (in our set of experiments, stimuli that were perceived) and D_2 for items from Source B (here, imagined stimuli). For items that are not successfully recognized as old or new, an old–new judgment is made randomly, influenced by bias b. Next, successfully detected items are subjected to source (in this case, reality monitoring) discrimination. d_1 and d_2 give the probability of successful source identification for Source A (in this case, perceived) and Source B (imagined) items, respectively. The bias parameter a gives the probability that a detected (i.e., recognized as old) but nondiscriminated (as to source) item will be labeled as originating from Source A. The bias parameter g gives the probability that a nondetected item will be labeled as belonging to Source A. Variants of this processing-tree model can be used to analyze 3 × 3 matrices constructed from the number of responses of each variety (Source A, Source B, and New) made to each of the three types of items. These analyses are performed by a software package by Hu

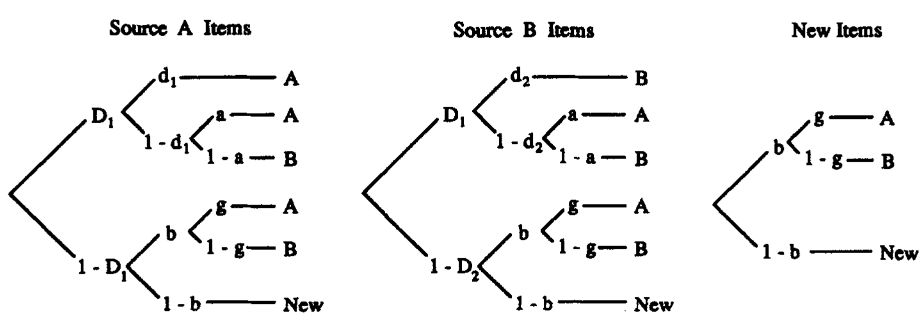

Figure A1. The seven-parameter multinomial model of source monitoring. D_1 = probability of detecting Source A (perceived) items as old; D_2 = probability of detecting Source B (imagined) items as old; d_1 = probability of discriminating the source of detected Source A (perceived) items; d_2 = probability of discriminating the source of detected Source B (imagined) items; b = bias for responding "old" to a nondetected item; a = probability of guessing that a detected but nondiscriminated item belongs to Source A; g = probability of guessing that a nondetected item belongs to Source A. Adapted from "Multinomial Processing Models of Source Monitoring" by W. H. Batchelder and D. M. Riefer, 1990, *Psychological Review, 97,* p. 551. Copyright 1990 by the American Psychological Association. Reprinted by permission of the author.

(1990). For theoretical and computational details concerning the software and the application of this approach see also Batchelder and Riefer.

The first step in applying the multinomial model analyses is to pick a particular model on which to base estimates of recognition and source-monitoring parameters. The overall seven-parameter model previously described (and in Figure A1) is, unfortunately, mathematically nonidentifiable (Bishop, Fienberg, & Holland, 1975; Greeno & Steiner, 1964). However, certain simpler models are identifiable, given particular assumptions. For instance, some models (viz., 4, 5a, 5b, and 6a, according to Batchelder and Riefer's (1990) nomenclature) assume that old–new recognition accuracy is the same for items from the two sources (perceived and imagined). In addition, some models (viz., 4, 5a, 5c, and 6b) assume that source-monitoring accuracy is identical for items from the two sources. The first of these two hypotheses (viz., identical recognition accuracies across sources) can be tested in a model-free fashion by comparing the proportion of perceived stimuli labeled as new with the proportion of imagined stimuli labeled as new. This was done separately for each response–signal lag condition in each of the experiments described here. The hypothesis that the proportion of perceived items labeled as new was equal to the proportion of imagined items labeled as new could not be rejected for any of these cases ($-.60 < Z < 1.25$). In addition, analyses using models that do assume different recognition accuracies across sources yielded virtually identical estimates of these recognition accuracies anyway. The analyses presented in this article were confined to models yielding a single-parameter estimate (D) for old–new recognition.

The next step is to decide whether to select a model that assumes different source- (in this case, reality) monitoring accuracies for perceived and imagined items or to use a model that yields a single-parameter estimate for source monitoring. For theoretical reasons, the question of different source discrimination accuracies for perceived and imagined items was of specific interest, so Model 5b was selected for all of the multinomial analyses reported here because it is the only (identifiable) model left yielding separate source-monitoring parameter estimates for perceived and imagined items. As Batchelder and Riefer (1990) noted, however, the assumptions that permit the identifiability of Model 5b (viz., where $D_1 = D_2$ and $a = g$ are assumed and where hypotheses of the form $d_1 = d_2$ can be tested) are not necessarily more appropriate for a given data set than those included in Model 5a (where $D_1 = D_2$ and $d_1 = d_2$ are assumed and hypotheses of the form $a = g$ can be tested). Unfortunately, there is no way to test the $a = g$ assumption with the current data set as there was for the $D_1 = D_2$ assumption. For this reason, an alternative to the interpretation of the results from Experiments 1 and 2 is that source discrimination is actually equivalent for perceived and imagined items, but that the bias for labeling a stimulus as perceived depends on whether it has been successfully detected (i.e., recognized). One solution to this problem of interchangeable models is to include three old sources, rather than two, in the experimental design. Riefer, Hu, and Batchelder (1994) used three sources in a study investigating memories for pictures versus words and did not find any difference in a and g that would confound conclusions about their obtained differences in source identification of pictures and words. Their task was similar but not identical to that used in our Experiment 1 in the response time condition. (Riefer et al.'s subjects were not explicitly instructed to generate images on word trials and were not under instructions to respond quickly on test trials.) Perhaps more critical for understanding and characterizing source monitoring, the alternative interpretation of the data from our experiments based on Model 5a would be quite inconsistent with the large body of previous research (e.g., Durso & Johnson, 1980; Finke et al., 1988; Intraub & Hoffman, 1992; Johnson et al., 1988; Lindsay et al., 1991; Rabinowitz, 1989), illustrating the importance of different kinds of evidence for the identification of memories of perceived and imagined events. For these reasons, Model 5b seems clearly preferable to Model 5a for describing the present data.

Received July 6, 1992
Revision received January 12, 1994
Accepted January 13, 1994 ■