

# Memory for script actions: Effects of relevance and detail expectancy

RUTH H. MAKI

*North Dakota State University, Fargo, North Dakota*

Script actions that varied both in relevance to the script and in the expectancy of their details were developed. In Experiment 1, recall of these actions was tested. Actions of medium relevance were recalled better than actions of high or low relevance, whether or not details were presented. However, the unexpected details themselves were recalled better than the expected details. In Experiment 2, recognition was lowest for high-relevance actions but medium- and low-relevance actions did not differ. Actions with unexpected details were recognized better than actions with expected details, which tended to be better than generic actions. The results were interpreted as showing that both the relevance and the expectancy of the details in script actions are important in memory, but that the two dimensions may play different roles.

Many studies have shown that prior knowledge affects the processing of new material—that schemata are used to comprehend and organize integrated material (Mandler, 1984; Taylor & Crocker, 1981). Generally, information important to an activated schema is remembered better than information less important (Johnson, 1970); however, other findings appear to conflict with this generalization. Using scripts, which are highly schematized sets of actions, Graesser, Woll, Kowalski, and Smith (1980) found that actions that were highly typical, and presumably important, were recalled and recognized more poorly than actions that were less typical. In person memory experiments, Hastie (1980) and Srull (1981; Srull, Lichtenstein, & Rothbart, 1985; Srull & Wyer, 1989) found that behaviors incongruent with trait expectations were recalled better than behaviors that were congruent. Thus, the literature on importance indicates that the more related material is to a schema, the better it will be remembered; but the script and person memory literatures seem to show the opposite.

The purpose in the present experiments was to understand when relatedness to a schema helps memory and when it hurts memory. Relatedness was manipulated in two ways. First, relevance of actions to an activated script was varied with low-, medium-, and high-relevance actions. The main question for this variable was whether relevance would produce an inverted U-shaped function in recall and a negative linear function in recognition. Second, the expectancy of the details contained in those actions was varied. The main question here was whether both recall and recognition would be better for unexpected

than for expected details.<sup>1</sup> Different patterns were predicted for the two types of relatedness, because it was assumed that script actions are at a higher level in a hierarchy of text propositions and that details are at a lower level. Patterns of recall and recognition were predicted from prior research and from three models: the schema-copy-plus-tag model (Graesser & Nakamura, 1982), the associative model of person memory (Srull, 1981), and the relatedness-distinctiveness model (Hunt & Einstein, 1981).

## A Hierarchical View of Actions and Details

Although the script and person memory findings are similar, they may involve manipulations at different levels, in the sense of the Kintsch's (1988) model of discourse processing. In research with scripts, manipulations of the relevance of script actions involve varying the degree of relationship of a script action to the overall goals of the script. Thus, the script actions may be represented as high-level macroproposition nodes. However, congruence of behaviors given an expectation of a personality characteristic may involve a lower level node—for example, a person is known to be friendly, but the specific behaviors give the details to support this. Behaviors in person memory experiments may be more similar to lower level microproposition nodes, or they may even be arguments filled into the slots of concept nodes. Thus, although better recall of less relevant script actions and better recall of incongruent behaviors appear to be similar phenomena, they may actually occur for different reasons, because they represent different levels in an associative representation of the text.

Details filled into scripts are more like arguments filled into slots than they are like concept nodes. Such details may function like instances in categories, whereas the actions in scripts probably function like schemata. Wyer and Gordon (1984) described some different characteristics of these two types of memory representations. Schemata

---

Appreciation is expressed to Kim Veum, Heidi Goos, Tom Tigue, and Nate Emerson for testing the subjects and scoring the data. Portions of this paper were presented at the annual meeting of the Psychonomic Society, Seattle, 1987. Requests for reprints should be sent to Ruth H. Maki, Department of Psychology, University Station—Box 5075, North Dakota State University, Fargo, ND 58105.

are more highly organized than categories. Relative to categories, the relevant actions of a schema are likely to be activated when a part of the schema is presented, so subjects are more likely to infer nonpresented actions (Barsalou & Sewell, 1985), and, therefore, intrusions are more likely with schemata. Subjects are less likely to infer nonpresented category instances, and, in the case of scripts, they may not fill in empty argument slots. The proposed relevance dimension is a schematic dimension, but the detail expectancy dimension may consist of category instances. If script actions and the details in argument slots really are like schemata and categories, then different patterns of recall and recognition might be observed when the likelihood of script actions and details is varied.

### Models of Relatedness and Memory

Situations in which less relevant material is remembered better than more relevant material are of particular interest for the present experiments. Several models have been proposed to explain such cases. The schema-copy-plus-tag model, which was developed from research with scripts, hypothesizes that subjects copy a portion of the information in a schema into a memory representation of the event (Graesser & Nakamura, 1982). The highly relevant information is copied. Material that actually has been presented cannot be discriminated from nonpresented material that is part of the copied schema. Thus, highly relevant actions will be recognized poorly because many false alarms will be made, and they will be recalled poorly because subjects cannot tell which actions were and were not presented. Less relevant material, however, is placed in memory with specific tags. Tagged irrelevant actions do not produce discrimination problems, and they are recognized and recalled well. Thus, the schema-copy-plus-tag model makes the clear prediction that low-relevance actions will be recalled better than high-relevance actions. The model does not make any specific predictions for medium-relevance actions. If they are part of the copied schema, they will be remembered more poorly than low-relevance actions, but if they are tagged, they will be remembered better than the copied high-relevance actions.

The schema-copy-plus-tag model also does not make predictions about details within scripts, because it was not formulated to account for memory for details. Still, it could be extended to predict such detail recall. Expected details in high-relevance actions might be part of the copied schema. Thus, they may be recognized and recalled poorly, just as high-relevance actions should be. Unexpected details in high-relevance actions, however, should be tagged and remembered well. The details in medium- and low-relevance actions should also be tagged and, therefore, remembered well. The schema-copy-plus-tag model, then, predicts an interaction between relevance and detail expectancy. Expected details in high-relevance actions should be remembered more poorly than all other details.

The associative model of person memory (Hastie, 1980; Srull, 1981) was developed to explain why incongruent

behaviors are recalled better than congruent behaviors. Hastie (1980) found that if subjects were given a trait expectancy about a person and then shown a list of that person's behaviors, recall would be highest for behaviors that were incongruent with the expectancy and lowest for behaviors that were irrelevant, with congruent behaviors in the middle. Srull's (1981) associative model for these results postulates that behaviors incongruent with expectancies are processed more in order to understand why an individual would perform them. This extended processing results in incongruent behaviors' being more highly associated both with other behaviors and with the person concept. Therefore, recall of incongruent behaviors is higher than recall of congruent behaviors. Further support for this model can be seen in a study by Hemsley and Marmurek (1982), who showed that incongruent behaviors were studied longer than congruent ones in conditions that produced higher recall of incongruent behaviors. Although Srull and Wyer (1989) have recently specified the situations in which incongruent behaviors will be recalled better and the situations in which they will not, the newer model still postulates better recall of material that is incongruent with expectations.

Extension of the associative model to memory for scripts yields clear predictions for detail expectancy. Unexpected details should be recalled better than expected details. They are incongruent with expectations, and they should be given more processing. However, whether this model also applies to the relatedness of actions is unclear. In the person memory research, irrelevant behaviors have been recalled more poorly than either congruent or incongruent behaviors, because they are not associated with other relevant behaviors. Thus, this model seems to predict poorer memory for irrelevant actions than for relevant ones. The associative model predicts no differences in recognition, because differences are due to retrieval, which plays less of a role in recognition than in recall.

The third model that is related to action relevance and detail expectancy is the relatedness-distinctiveness model developed by Hunt and his colleagues (Hunt & Einstein, 1981; Hunt, Ausley, & Schulz, 1986). They distinguish between the relational, organizational aspects of materials, which suggest, for example, how well a set of items can be integrated, and item-specific information, or "distinctiveness" (see esp. Hunt et al., 1986). Although this theory was developed for word lists, it has been extended to memory for prose (Einstein, McDaniel, Bowers, & Stevens, 1984). The present relevance dimension probably manipulates the relational properties of the actions. The most relevant actions would be the most highly organized with respect to the script, and the least relevant actions would be less well organized; but highly relevant actions are also the least distinctive, so their recall would be lowered. Recall of low-relevance actions should also be low, because their retrieval is unlikely with the script name as a cue. Medium-relevance items would be recalled best; they are moderately related, likely therefore to be retrieved, and distinctive. Thus, the relatedness-distinctiveness model (Hunt & Einstein, 1981) predicts an inverted

U-shaped function between action recall and relevance. The expectancy of details would manipulate the distinctiveness of the actions. Given equal relatedness, the more distinctive unexpected details should be remembered better than the less distinctive expected details.

According to the relatedness-distinctiveness model, retrieval is less important in recognition than in recall, so distinctiveness plays a much greater role than relatedness. Therefore, in recognition, there should be a negative relationship between memory and relevance, and unexpected details should be recognized better than expected ones.

### EXPERIMENT 1

Recall was tested in Experiment 1 for high-, medium-, and low-relevance actions that contained details either expected or unexpected. Three questions were asked in Experiment 1. The first one involved the relationship between recall and relevance, the second asked about recall of expected as compared to unexpected details, and the third asked whether action relevance and detail expectancy produced the same patterns in recall or different patterns.

An inverted U-shaped function relating relevance to recall was predicted, for two reasons. First, prior research suggested that such a relationship exists. Script experiments (e.g., Graesser et al., 1980) show that highly relevant actions are recalled poorly. Script experiments do not provide evidence about the recall of medium- and low-relevance actions, because atypical actions in those experiments have usually included a combination of what I have defined as medium- and low-relevance actions (Graesser et al., 1980). However, experiments manipulating importance with narrative or expository prose have usually shown that more important material is recalled better than less important material (Johnson, 1970; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). The prose used in importance experiments does not usually include very obvious, totally predictable ideas. Thus, the level of relevance of important ideas in prose experiments probably corresponds roughly with medium-relevance actions in scripts. So, medium-relevance (more important) actions should be recalled better than low-relevance (less important) actions. Therefore, script experiments lead to the prediction that high-relevance actions will be recalled poorly, and importance experiments lead to the prediction that medium-relevance actions will be recalled better than low-relevance actions. Overall, then, recall of medium-relevance actions should be higher than recall of high- and low-relevance actions.

The second reason for the prediction of the inverted U-shaped relationship comes from the relatedness-distinctiveness model (Hunt & Einstein, 1981), which argues that both retrievability and distinctiveness determine recall. The medium-relevance actions, which are retrievable yet distinctive, should be remembered better than high-relevance actions, which are not distinctive, and low-relevance actions, which are not retrievable.

Comparing high- and low-relevance actions is also of interest, because the schema-copy-plus-tag model (Graesser & Nakamura, 1982) predicts that low-relevance actions will be remembered better than high-relevance actions. If the associative model is extended to the action level, then, on the basis of its predictions of poorer recall of irrelevant as opposed to relevant behaviors, it will predict poorer recall of low-relevance as opposed to high-relevance actions.

The next question addressed in Experiment 1 involved the role of detail expectancy. As stated earlier, both the associative and the relatedness-distinctiveness models predict better recall of unexpected details. The schema-copy-plus-tag model predicts an interaction between detail expectancy and relevance. For highly relevant actions, unexpected details should be tagged and remembered well, but expected details may be part of the copied schema. Recall of expected and unexpected details in medium- and low-relevance actions should not differ, because neither would be a part of the copied schema; all would be tagged and remembered well.

The third question addressed in Experiment 1 was whether relevance and detail expectancy were both manipulating the same dimension or whether they actually are different dimensions. This question was investigated by comparing the patterns of results for the two variables. There has not been much research in which relevance and the expectancy of details have both been manipulated. In one of the few such experiments, Bellezza and Bower (1982) presented script actions that were relevant with expected objects (e.g., "She finally was able to find a seat" in an "Attends a Lecture" script), relevant with unexpected objects (e.g., "She then opened up her deck of cards" instead of "opened up her notebook" in the lecture script), or irrelevant. Subjects rated the relevant actions with expected objects as more likely than relevant actions with unexpected objects and more likely than irrelevant actions. In spite of the large difference in rated likelihood, recall of relevant actions with expected and with unexpected objects was identical and higher than recall of irrelevant actions. Bellezza (1983) reported that the pattern of study times matched the pattern of likelihood ratings. Thus, both likelihood ratings and study times were similar for relevant actions with unexpected objects and for irrelevant actions, yet the recall of relevant actions with unexpected objects was much higher. Their recall may have been boosted by the unexpectedness of their details.

One difficulty in interpreting the Bellezza experiments (Bellezza, 1983; Bellezza & Bower, 1982) is that the manipulation of expected and unexpected objects also manipulated the relevance of the script actions. That is, "opening up a deck of cards" makes that action irrelevant to the lecture script, whereas "opening up a notebook" makes that action relevant. Thus, it is difficult to tell whether relevance and detail expectancy affected recall in similar or different ways. In Experiment 1, care was

**Table 1**  
**Examples of Actions Varying in Relevance and Details**

Details	Relevance		
	High	Medium	Low
Generic	He got his bike out	He cleaned off his bike	He stretched
	She put some food on to cook	She decided when to serve dinner	She fed the dog
Expected	He got his bike out of the garage	He cleaned off the seat of his bike	He stretched his legs
	She put some food in the oven to cook	She decided to serve dinner at about 6 p.m.	She fed the dog his dog food
Unexpected	He got his bike out of the kitchen	He cleaned off the spokes of his bike	He stretched his arms
	She put some food in the toaster to cook	She decided to serve dinner at about 10 p.m.	She fed the dog his science diet

taken to manipulate relevance and detail expectancy as independently as possible, so that their roles in memory could be compared.

### Method

**Materials.** Script actions that varied both in relevance to the script and in the expectancy of details were developed. For the relevance dimension, generation and rating procedures similar to those developed by Graesser, Gordon, and Sawyer (1979) were used. Twelve subjects generated script titles; the 20 most frequent titles were selected. A group of 39 new subjects generated script actions. The frequency with which each action was generated was counted independently by two scorers. Actions were considered to be high frequency if both raters agreed that the specific action was generated by at least 30% of the subjects. Actions that were generated by only 1 subject and actions that were not generated by any subject were used as low-frequency actions. The actions from each script were put together in temporal order, and these organized actions were used in other rating and generation tasks. A group of 27 subjects rated high- and low-frequency actions on a 7-point scale, with 1 = *very unimportant* in an average enactment of the script and 7 = *very important*. The subjects were asked to give importance rather than relevance ratings, because the experimenters felt that importance would be understood better by the subjects. Actions that were generated by 30% of the subjects and given a mean rating between 5.4 and 7 were classified as high-relevance actions; actions generated by 1 or no subjects and given mean ratings between 3.8 and 5.3 were considered to be of medium relevance; actions given ratings below 3.5 were low-relevance actions. Thus, relevance to the script was defined in terms of both generation frequency and importance ratings. The 10 scripts that had at least eight actions in each of these relevance categories were selected for detail generation.

Most of the script actions generated by the subjects were generic (e.g., "decide what to fix" in the "making a meal script"). A question designed to fill in a detail was framed for each script action. For example, the question "What?" was used for "decide what to fix." A group of 29 subjects gave details in answer to the questions "What?" (for 46% of the actions), "Where?" (15%), "Why?" (13%), "How?" (11%), "When?" (11%), and "Who?" (6%). Expected details were those given by at least 20% of the subjects. Unexpected details were either given by only 1 subject or made up by the scorers. Next, the expectancy of each detail was rated by 24 subjects. The rating sheets were similar to those used in the detail generation phase, with the action, a question in brackets, and a detail. Subjects used a 7-point scale ranging from 1 = *not at all expected* to 7 = *very expected* to rate the expectancy of the

detail. The mean expectancy rating for all the actions in each script was calculated. Actions with details that were rated one half of a standard deviation above the mean or higher were considered expected, and actions that were rated one half of a standard deviation below the mean or lower were considered unexpected. Finally, generic actions and those with expected and unexpected details were rated for importance by 44 subjects, to determine whether adding details affected perceived importance. This rating scale ranged from 1 = *easily eliminated* to 7 = *very important*.

Six scripts that contained at least 8 expected and 8 unexpected actions at each level of relevance were selected. These six scripts were "Going on vacation by car," "Going for a bike ride," "Preparing an assignment for class," "Making a meal at home," "Writing a letter," and "Going to the lakes." Examples of the various types of actions for the bike ride and the meal scripts are shown in Table 1.

Mean expectancy ratings and importance ratings for these actions as a function of relevance and details are shown in Table 2. Subscripts indicate which means are not different. As can be seen in the left half of the table, expectancy ratings were higher for expected than for unexpected details. In addition, details in more relevant actions were rated as somewhat more expected than details in less relevant actions. The mean importance ratings for these actions are shown in the right half of Table 2. Importance ratings increased with greater relevance, but relevance and details interacted. The addition of unexpected details reduced the perceived importance of actions. The addition of expected details reduced the perceived importance of highly relevant actions, but expected details did not affect the perceived importance of medium- and low-relevance actions. Because detail expectancy ratings were affected by relevance, and because importance ratings were affected by details, these two dimensions are not independent. However, the differences in the manipulated dimension were much greater than the differences in the other dimension for both relevance and detail ex-

**Table 2**  
**Mean Detail Expectancy and Relevance Ratings of Actions Used in the Experimental Scripts**

Relevance	Detail Expectancy Ratings		Relevance Ratings		
	Expected	Unexpected	Generic	Expected	Unexpected
High	5.90 <sup>a</sup>	2.48 <sup>c</sup>	6.38	5.67	4.50
Medium	5.64 <sup>ab</sup>	2.28 <sup>cd</sup>	4.81 <sup>a</sup>	4.79 <sup>a</sup>	3.48
Low	5.18 <sup>b</sup>	2.24 <sup>d</sup>	2.84 <sup>b</sup>	3.01 <sup>b</sup>	2.22

Note—Within each type of rating, means with shared superscripts are not significantly different.

pectancy. The large within-dimension differences should play a greater role in memory than the smaller between-dimension differences.

Six versions of each script were created. There was a total of 24 actions for each script; half the subjects saw scripts containing 12 of the actions and the other half saw scripts containing the other 12 actions. Each script version contained 12 critical actions plus 3 filler actions to improve its flow. Four of the script versions contained details, and two were generic scripts without details. There were 4 actions at each relevance level, and for the scripts with details, 2 of these actions contained expected details and 2 contained unexpected details. Specific actions, generic versus detailed scripts, and expected versus unexpected details were counterbalanced across subjects. Two other scripts ("Attending a lecture in a class," "Going grocery shopping"), which did not have 24 actions that met the relevance and expectancy criteria, were used as primacy and recency buffers; the data from these scripts were not analyzed. For half the subjects, the primacy script contained details and the recency script was generic; this was reversed for the other half of the subjects.

**Procedure.** The subjects were informed that they would see several stories and that they should read the stories carefully because they would be asked some questions about them later. Script titles and actions were presented individually on a monitor attached to an Apple II+ computer. The title of each script was presented in capital letters followed by each action in chronological order. The subjects pressed the space bar on the computer to erase each item and to present the next one. The reading time for each item was recorded. After one script was finished, the computer produced two beeps followed by the title of the next script. After all eight scripts had been read in this way, the subjects performed a number progression filler task for 10 min. The subjects were to fill in the last numbers in sequences of numbers arranged according to different rules. Then they began the recall task.

In the recall task, the subjects were given eight sheets of paper with different script titles at the top. The titles were arranged in the order in which the scripts had been presented. The subjects were instructed to write everything that they could remember from each script. They were told that they did not need to try to write the exact words that were used, but that they should try to write down all of the ideas as completely as possible. They were given as much time as they needed to recall all that they could from each script.

**Subjects.** Eighteen volunteer subjects from introductory psychology courses at North Dakota State University participated for extra credit toward their course grades.<sup>2</sup>

## Results and Discussion

**Scoring of recall.** Each script was scored by two independent scorers, using a 3-point scale. Actions were scored first. If the subject recalled the action as a close paraphrase of the presented action, 2 points were given. If the action was recalled but was less similar to the original, 1 point was given. Zero was given if the action was not mentioned. The recall of actions was scored the same, whether they had been presented in generic or detailed scripts. Interrater reliability for the two scorers, as measured by the median Pearson  $r$ , was .88. Scripts with details were also scored for the presence of details. If a detail was given accurately, 2 points were given; if a detail was mentioned but was less similar to the original, 1 point was given; if a detail was omitted, a 0 was assigned. The median Pearson  $r$  for interrater reliability for the details was .90. Intrusions were scored in the same manner. If a subject wrote an action from the set that was not presented, or if a subject added a detail from the other set,

it was scored as an intrusion. Thus, extra items were only counted as intrusions if they came from the nonpresented set, making the sizes of the presented and nonpresented sets of actions the same. The average score given by the two raters to each item was used in the recall analyses.

**Recall of actions.** Recall of actions, whether or not details had been presented, was analyzed first. Across scripts, there was a total of 8 actions in each condition; the maximum recall score for each action was 2, so the total score possible in each condition was 16. Recall scores are shown in Table 3. Because it is important to correct recall for guessing when highly relevant material is presented, the main analysis was conducted on recall scores corrected for intrusions. The correction that was used was the memory improvement score (MI) that Graesser et al. (1980) used for similar data:  $[p(\text{recall}) - p(\text{intrusion})] / [1 - p(\text{intrusion})]$ . The MI scores for action recall are presented in Table 3.

The MI scores were analyzed, first, with subjects as the random variable, and then with script action sets as the random variable. For the subjects analysis, the design was a 3 (relevance)  $\times$  3 (detail expectancy) within-subjects analysis of variance (ANOVA). Whether actions were generic or whether they contained expected or unexpected details did not affect recall of the actions themselves ( $F < 1$ ), so this variable was averaged across in Table 3. A planned comparison showed that the score for medium-relevance actions (.29) was higher than for the average of high- and low-relevance actions combined [.24;  $F(1,17) = 7.39$ ,  $MS_e = .01$ ]. A second planned comparison indicated that the recall of high and low actions did not differ ( $F < 1$ ).

Next, the data were analyzed with scripts and action sets as the random variable. There were six scripts having two sets of actions each. Mean MI scores were calculated for each of these 12 item sets. An ANOVA on these scores showed that the planned comparison combining high- and low-versus medium-relevance actions was significant [ $F(1,11) = 5.05$ ,  $MS_e = .12$ ]. The second comparison, between high- and low-relevance actions, was not significant across items ( $F < 1$ ). Thus, medium-relevance actions were recalled better than high- and low-relevance actions, whether the analysis was done with subjects or with script action sets as the random variable.

The prediction that medium-relevance actions would be recalled best was supported. The fact that high- and low-relevance actions did not differ contrasts with results found

Table 3  
Mean Memory Improvement (MI) and Recall Scores  
for Script Actions and Details

Type	Relevance					
	High		Medium		Low	
	MI	Recall	MI	Recall	MI	Recall
Actions	.23	4.15	.29	4.69	.24	3.87
Details						
Expected	.17	2.94	.20	3.17	.15	2.64
Unexpected	.25	4.03	.29	4.71	.26	4.08

by Graesser (Graesser et al., 1980; Smith & Graesser, 1981), who reported that low-relevance actions were recalled somewhat better than high-relevance actions, especially with a correction for intrusions. However, Graesser's low-relevance actions were a mix of what we have defined here as medium- and low-relevance actions. The medium-relevance actions here produced the highest recall, so if they were mixed with low-relevance actions, they might produce better recall than high-relevance actions would.

**Recall of details.** Detail recall was scored in the expected and unexpected detail conditions in the same manner as was action recall. Although intrusions were fairly rare, the MI score used by Graesser et al. (1980) was also computed for detail recall. These corrected recall scores are presented in Table 3, along with the raw scores (out of a possible 16).

The MI scores for details were also analyzed in one ANOVA with subjects as the random variable and in another ANOVA with script action sets as the random variable. The ANOVA with subjects indicated that unexpected details were recalled better than expected details [ $F(1,17) = 7.85$ ,  $MS_e = .15$ ]. The average recall score of details in medium-relevance actions (.24) was somewhat higher than the recall score of details in high- and low-relevance actions (.21), but this effect was not significant [ $F(1,17) = 2.46$ ,  $MS_e = .01$ ]. The recall of details in high- and low-relevance actions did not differ ( $F < 1$ ).

Similar effects were seen across script action sets. Unexpected details were recalled better than expected details [ $F(1,11) = 13.62$ ,  $MS_e = .01$ ]. Details in medium-relevance actions tended to be recalled better than details in high- and low-relevance actions, but this effect was not significant [ $F(1,11) = 2.80$ ,  $MS_e = .006$ ]. The recall of details in high- and low-relevance actions did not differ ( $F < 1$ ).

Unexpected details were recalled better than expected details. This finding may be similar to the congruency effect in person memory experiments (Hastie, 1980; Srull, 1981; Srull et al., 1985): Subjects remember behaviors that are incongruent with expectations better than they remember behaviors that are congruent. The incongruency effect presumably results from more processing of the incongruent behaviors, which are studied longer (Hemsley & Marmurek, 1982). If the present effect is similar to this person memory effect, then longer study times should occur for unexpected than for expected details. The study time data are presented following Experiment 2, so that study times for both experiments can be combined. Relevance produced the same pattern of effects for detail recall as for action recall, but the differences were small and not significant when details, as opposed to action recall, were considered.

In Experiment 1, recall showed the U-shaped pattern predicted by the relatedness-distinctiveness model (Hunt & Einstein, 1981), but not predicted by either the schema-copy-plus-tag model (Graesser & Nakamura, 1982) or the

associative model (Srull, 1981). The difference in high- and low-relevance actions predicted by the schema-copy-plus-tag model did not occur. Experiment 1 also showed that unexpected details were recalled better than expected details, as was predicted by both the associative model and the relatedness-distinctiveness model. In contrast to the prediction from the schema-copy-plus-tag model, there was no interaction between relevance and detail expectancy in detail recall. Thus, on the surface, the recall results appear to support the more general relatedness-distinctiveness model better than either the schema-copy-plus-tag model or the associative model, although either of the latter models could be modified to account for the observed effects. Experiment 2 was designed to test predictions by the three models in a recognition task.

## EXPERIMENT 2

The three models make different predictions for recognition than they do for recall. Both the schema-copy-plus-tag (Graesser & Nakamura, 1982) and the relatedness-distinctiveness model (Hunt & Einstein, 1981) predict better recognition of lower relevance script actions. The schema-copy-plus-tag model makes this prediction because the low-relevance actions are more likely to be tagged than the high-relevance actions are. If both medium- and low-relevance actions are tagged, they should not differ in recognition, but both should be better than high-relevance actions. In other script experiments, poorer recognition of high-relevance actions has occurred, because there is a very high false alarm rate for relevant distractors (Graesser et al., 1980); hit rates usually do not vary with relevance. The basic problem in recognizing highly relevant actions is that subjects are unable to discriminate which ones actually were presented from those that could have been presented, but were not. A similar prediction is made by Hunt and Einstein's (1981) model because lower relevance actions should be more distinctive, but this model predicts a difference in hits as well as in false alarms.

As with recall, the schema-copy-plus-tag model predicts an interaction between relevance and detail expectancy. Tagged unexpected details should be recognized better than copied expected details in high-relevance actions, but expected and unexpected details should not differ in medium- and low-relevance actions. Hunt and Einstein's (1981) model clearly predicts better recognition of unexpected details; they are more distinctive than expected details are. Srull's (1981) version of the associative model does not predict effects in recognition of either relevance or detail expectancy, because the congruency effects are due to retrieval, which is minimized in recognition. More recently, however, Wyer, Bodenhausen, and Srull (1984) found the reverse of a congruency effect in recognition: Congruent behaviors were recognized better than incongruent ones, although this effect was found only for impression formation and not for memory instructions. In a similar recognition experiment, Marmurek (1984)

manipulated the proportions of congruent and incongruent behaviors in a list and found that recognition was higher for the behaviors that were in the minority. Thus, Marmurek concluded that behaviors inconsistent with the final impression were recognized better. Marmurek's effect occurred mostly in the hit data: The false alarm rate was fairly low (about .10).

The main difference between these person memory experiments and the script experiments appears to be in the false alarm rates. Highly relevant script actions are recognized poorly because false alarms in response to non-presented items are high; incongruent and congruent behaviors do not consistently produce differential rates of false alarms. If unexpected details are at the same level as behaviors in person memory experiments, any difference in recognition of expected and unexpected details may occur in hits and not in false alarms. Differences as a function of relevance should occur because of differences in false alarms and not hits, as Graesser et al. (1980) have found. The distinctiveness theory appears to predict a difference in both hits and false alarms.

## Method

**Materials.** The six critical and two buffer scripts from Experiment 1 were used. Again, one of the buffer scripts contained details and one was generic; four of the critical scripts contained details and two were generic scripts. Specific actions and details were counterbalanced across script versions, as in Experiment 1.

**Procedure.** The presentation procedures were identical to those of Experiment 1. Half the subjects took the recognition test immediately, and half took it following a 1-week delay. On the recognition test, each script title was presented followed by 24 script actions, 12 of which had been seen by each subject. The other 12 distractor actions were of the same type (generic or half with expected and half with unexpected details) as the old, presented actions. The subjects were instructed to press the slash key on the right of the computer keyboard if the action had been presented when they read the script and to press the "Z" key on the left of the keyboard if it had not. The key instructions were printed below each action, at the bottom of the screen. After each subject had pressed the appropriate key, a rating scale appeared on the screen. It varied from 1 = *just guessing* to 7 = *very sure*. The subjects were instructed to type a number from 1 to 7 to indicate their confidence in the recognition decision that they had just made.<sup>3</sup> The scripts were tested in the same order as they had been presented, and the script actions were presented in their appropriate order, although the subject had only seen half of them before. The two buffer scripts were tested in the first and last positions, but the data from them were not used.

**Subjects.** Eighteen subjects were randomly assigned to the delay condition, and 18 subjects were randomly assigned to the im-

mediate condition. All subjects were volunteers from introductory psychology courses at North Dakota State University who were participating for extra credit.

## Results and Discussion

**Analysis of hits and false alarms.** Proportion hits and false alarms were calculated, both for the four scripts that contained expected and unexpected details and for the two generic scripts. Because the interest was in the specific patterns of hits and false alarms, these were analyzed separately. Proportion hits are shown in the top row of Table 4. These were analyzed in a 2 (delay)  $\times$  3 (relevance)  $\times$  3 (detail expectancy) mixed-design ANOVA, with subjects as the random variable. There were no significant effects in the hit data, including the difference between hits in the immediate (.78) and delay (.70) conditions [ $F(1,34) = 2.54$ ,  $MS_e = .20$ ,  $p > .05$ ]. This finding is similar to what Graesser and his colleagues (Graesser et al., 1979; Graesser et al., 1980; Smith & Graesser, 1981) have reported in their experiments with script actions.

More false alarms were made in the delay (.34) than in the immediate condition (.21;  $F(1,34) = 12.28$ ,  $MS_e = .12$ ). Delay and detail expectancy interacted, because the increase in false alarms from the immediate to the delay condition was greater in the generic (.31 to .49) and expected (.22 to .39) conditions than in the unexpected condition [.08 to .14;  $F(2,68) = 6.25$ ,  $MS_e = .02$ ]. Relevance and detail expectancy also interacted [ $F(4,136) = 10.58$ ,  $MS_e = .02$ ].

The relevance  $\times$  detail expectancy interaction was further analyzed as two separate 2  $\times$  3 factorials, with generic versus expected details and relevance in one analysis, and expected versus unexpected details and relevance in the second. In the generic versus expected analysis, details and relevance interacted [ $F(2,68) = 8.69$ ,  $MS_e = .03$ ]. For high-relevance actions, generic actions produced more false alarms than did actions with expected details [ $F(1,34) = 24.80$ ,  $MS_e = .03$ ], but generic and expected items did not differ at medium and low relevance [ $F_s(1,34) \leq 1.13$ ]. This pattern is similar to that of the importance ratings; generic actions were rated as more important than the same actions with expected details in the high-relevance condition. The highly relevant generic actions match the prototypical script closely, and thus subjects have difficulty discriminating which actions were and which were not presented. Adding details evidently makes them more discriminable.

Table 4  
Proportion Hits and False Alarms and  $d$ 's as a Function of Relevance and Details

	Relevance								
	High			Medium			Low		
	Generic	Expected	Unexpected	Generic	Expected	Unexpected	Generic	Expected	Unexpected
Hits	.77 <sup>a</sup>	.78 <sup>a</sup>	.72 <sup>a</sup>	.70 <sup>a</sup>	.74 <sup>a</sup>	.72 <sup>a</sup>	.68 <sup>a</sup>	.76 <sup>a</sup>	.76 <sup>a</sup>
False Alarms	.64	.40	.18	.33 <sup>b</sup>	.28 <sup>b</sup>	.07	.25 <sup>c</sup>	.23 <sup>c</sup>	.09
$d'$	.52	1.34	1.85	1.21	1.57	2.55	1.52	2.01	2.64

Note—Within each level of relevance, cell means with same superscripts are not significantly different.

In the analysis using only actions with expected and unexpected details, relevance and details did not interact [ $F(2,68) = 1.63$ ,  $MS_e = .01$ ]. However, actions with expected details produced more false alarms than actions with unexpected details [ $F(1,34) = 88.96$ ,  $MS_e = .02$ ]. The main effect of relevance was also significant [ $F(2,68) = 19.71$ ,  $MS_e = .02$ ]. Pairwise comparisons indicated that high-relevance actions produced more false alarms than did medium-relevance actions, but medium- and low-relevance actions did not differ.

Similar analyses were also conducted with script action sets as the random variable, and the results were essentially the same as with subjects as the random variable. For each subject,  $d$ 's were also calculated; their means are shown in Table 4. Their pattern was essentially the same as that for false alarm rates, and statistical analysis of them resulted in similar conclusions as with false alarms, except that there was no interaction between relevance and detail expectancy.

Actions with expected details produced more false alarms than those with unexpected details. Thus, details do function like category instances in recognition, with atypical instances producing better recognition than typical instances. High-relevance actions resulted in significantly poorer recognition than did medium- and low-relevance actions, but medium- and low-relevance actions did not differ significantly.

**Reading times.** The instructions for reading the scripts were the same in Experiments 1 and 2, so the data for reading times were combined across experiments. The reading time for each script action was divided by the number of syllables in the sentence, to correct for sentence length. The reading times per syllable were analyzed in a 2 (delay)  $\times$  3 (relevance)  $\times$  3 (detail expectancy) mixed ANOVA. Mean reading times per syllable in seconds can be seen in Table 5. There was a main effect of detail expectancy [ $F(2,136) = 13.30$ ,  $MS_e = .02$ ], a main effect of relevance [ $F(2,136) = 3.36$ ,  $MS_e = .01$ ], and an interaction between relevance and detail expectancy [ $F(4,272) = 2.47$ ,  $MS_e = .01$ ]. Pairwise comparisons indicated that generic actions were studied longer than were actions with both expected and unexpected details, which did not differ from each other. The interaction occurred because relevance did not produce an effect with actions containing expected details [ $F(2,136) = 2.11$ ,  $MS_e = .01$ ], or with actions containing unexpected details ( $F < 1$ ), but relevance was significant with generic actions [ $F(2,136) = 4.41$ ,  $MS_e = .02$ ]. Generic actions of medium relevance were studied longer than generic actions of high or low relevance, which did not differ.

**Table 5**  
Mean Reading Times per Syllable for Experiments 1 and 2 Combined

Details	Relevance		
	High	Medium	Low
Generic	.480 <sup>a</sup>	.512 <sup>b</sup>	.451 <sup>a</sup>
Expected	.396 <sup>c</sup>	.426 <sup>c</sup>	.429 <sup>c</sup>
Unexpected	.431 <sup>c</sup>	.440 <sup>c</sup>	.425 <sup>c</sup>

Note—Means with the same superscripts are not significantly different.

Generally, the relationship between study time and measures of memory was not orderly. Relevance had large effects on recall and recognition, but it only affected study time for generic actions. The pattern for generic actions—longer study times for actions of medium relevance than for actions of high and low relevance—did parallel the recall data, but the relevance effect for recall was found for all three detail conditions. Thus, higher recall of generic medium-relevance actions might be partially explained by longer study times, but, clearly, that is not the entire explanation, because the inverted U-shaped effect was more general for the recall data than for the study time data. The recognition data cannot be explained by study time at all; there was a negative linear relationship between recognition and relevance, which was not seen in study time. Higher recall of unexpected actions cannot be explained by study time either. Generic actions were studied the longest, possibly because subjects were filling in details. However, this did not result in better memory, as measured by either recall or recognition.

## GENERAL DISCUSSION

The relevance and expectancy ratings indicated that relevance and detail expectancy were not completely independent. The interaction in relevance ratings occurred because the ratings of generic actions and actions with expected and unexpected details all differed for highly relevant actions, but only the ratings of actions with unexpected details were lower for medium- and low-relevance actions. However, in recall of the generic actions themselves, detail expectancy produced no effects. Expectancy ratings differed as a function of relevance, because the details in highly relevant actions were rated as more expected than the details in low-relevance actions, but neither differed from medium-relevance actions. In contrast, relevance produced a small inverted U pattern in detail recall, not the decreasing linear one predicted from the expectancy ratings. In recognition, detail expectancy and relevance did interact, but only when generic actions were compared with expected details and not when expected and unexpected details were compared. Thus, overall, the patterns of memory were not predictable from the small interactions of relevance and detail expectancy in the relevance and expectancy ratings. The large differences in the relevance ratings when relevance was manipulated and in the expectancy ratings when detail expectancy was manipulated played larger roles in memory than the smaller differences that existed in nonmanipulated dimensions.

These two dimensions did produce different patterns in recall, and they did not interact. The function relating relevance and recall had an inverted U shape because medium-relevance actions were recalled better than high- and low-relevance actions. This pattern fits nicely with data from importance experiments which show that more relevant material is remembered better than less relevant material (e.g., Johnson, 1970), if it is assumed that material varies from medium to low relevance in importance



studies. Although the recall of high- and low-relevance actions did not differ, unexpected details were recalled better than expected details.

The patterns of recognition were similar for relevance and detail expectancy. Highly relevant actions produced more false alarms than medium- and low-relevance actions did. Actions with expected details produced more false alarms than actions with unexpected details did. Highly relevant generic actions produced more false alarms than did highly relevant actions with expected details. Overall, this pattern suggests that the more an action or detail overlapped with the prototypical script, the higher the false alarm rate.

These data will be discussed in terms of the three models described in the introduction: the schema-copy-plus-tag model (Graesser & Nakamura, 1982), the associative model developed for person memory experiments (Hastie, 1980; Srull, 1981; Wyer et al., 1984), and the Hunt and Einstein (1981) relatedness-distinctiveness model.

The basic problem for the schema-copy-plus-tag model is the positive role played by relevance in recall. The schema-copy-plus-tag model appears to predict that moderately relevant actions should be recalled more poorly than low-relevance actions, because they are more likely to be a part of the copied schema and subjects should have difficulty in editing presented and nonpresented actions. Alternatively, it might be argued that both medium- and low-relevance actions are equally likely to be tagged rather than copied from the schema; if so, they should be recalled equally well. This was not the case; medium-relevance actions were recalled better than low-relevance actions. Thus, the recall data are inconsistent with the schema-copy-plus-tag model. The recognition results are easily accounted for by this model. The more likely an action was to be a part of the copied schema (as for high-relevance actions), the more false alarms it produced.

The fact that unexpected details were recalled and recognized better than expected details at all levels of relevance also presents problems for the Graesser and Nakamura (1982) model. Whereas expected details might be part of the material that is copied from the prototypical script when they are part of high-relevance actions, they should not be part of the copied schema for medium- or low-relevance actions. If so, expected details would be recognized and recalled poorly in high-relevance but not in medium- or low-relevance actions; this was not the case. The schema-copy-plus-tag model simply was not developed to account for details filled into the action slots, so it would need to be extended to explain why unexpected details should be recalled and recognized better than expected ones, independently of relevance.

In contrast to the schema-copy-plus-tag model, the associative model of person memory was developed to account for recall of lower level propositions, which might be considered instances of categories in some cases or slots in scripts in the present case. It can easily explain the better recognition of unexpected details by assuming that they were processed more deeply than expected details, but it cannot handle the relevance effects. This model argues

that irrelevant behaviors are not remembered, because they are not associated with other behaviors. This explains the poor recall of low-relevance actions, but not the better recall of medium- rather than high-relevance actions. The recognition data are not handled by Srull's (1981) version of this model, because he located the effects in the retrieval process and predicted no recognition differences. However, another version of the model (Wyer et al., 1984) suggests that recognition is guided by familiarity. Generally, recognition should be higher for material that is more familiar in the context—that is, for congruent items. In fact, Wyer et al. (1984) found better recognition of congruent behaviors in most of their conditions, even when recognition was corrected for false alarms. In the present case, false alarms were higher for non-presented relevant actions and expected details, which have more familiarity in the script context, than they were for less relevant actions and unexpected details; but hits were not. Thus, familiarity in the script context increased errors but not correct recognitions.

An extension of the associative model might explain the present results in terms of informativeness and depth of processing. High- and low-relevance actions are not very informative; high-relevance, because they are obvious, and low-relevance, because they are not related to the script. Thus, medium-relevance actions would be processed most deeply because they are the most informative. Such actions should be recalled best, and they were. Unexpected details were also recalled better; they were more informative, in that expected details might be filled into generic actions almost automatically. However, the study time data did not support this interpretation. Medium-relevance actions were studied longer than high- and low-relevance actions only in the generic condition, but medium-relevance actions were recalled better in the two detail conditions also. Furthermore, actions with unexpected details were not studied longer than actions with expected details. Thus, overall, more highly informative material may have been recalled better, but it was not studied longer. In addition, the associative model cannot account for the recognition results even with a familiarity interpretation added.

Hunt and his colleagues (Hunt et al., 1986; Hunt & Einstein, 1981) distinguish between the relational or organizational property of materials (i.e., how well can a set of items be integrated) and item-specific information, or what Hunt et al. (1986) call "distinctiveness." This model predicts an inverted U-shaped function relating relevance and recall. Highly relevant actions should have been retrieved; but they were not distinctive in memory, and therefore they were not recalled well. Low-relevance actions would not be retrieved. Medium-relevance actions, however, would have a moderate probability of being retrieved with the script title as a cue, and they would be distinctive enough to be recalled.

This two-process model can also account for the recognition results, insofar as it suggests that recognition depends much more on distinctiveness than on relational properties of materials. The more distinctive an action

was, the better it was discriminated from nonpresented actions. Thus, medium- and low-relevance actions were recognized better than high-relevance actions, and actions with unexpected details were recognized better than actions with expected details, which were better than the least distinctive generic actions. However, this effect occurred only in false alarms and not in hits—a result predicted by the schema-copy-plus-tag model, but not by the relatedness-distinctiveness model.

Overall, the model proposed by Hunt and colleagues (Hunt et al., 1986; Hunt & Einstein, 1981) seems to account for the data best, perhaps because it is the least specific. However, it does not explain how items become distinctive or how the relational properties are defined for a given set of materials. The associative model may help one understand how distinctiveness develops, and the schema-copy-plus-tag model may help to explain which actions will be highly related to script titles. Thus, each of the models discussed above may have something to contribute to an understanding of the present results.

#### REFERENCES

- BARSALOU, L. W., & SEWELL, D. R. (1985). Contrasting the representation of scripts and categories. *Journal of Memory & Language*, *24*, 646-665.
- BELLEZZA, F. S. (1983). Recalling script-based text: The role of selective processing and schematic cues. *Bulletin of the Psychonomic Society*, *21*, 267-270.
- BELLEZZA, F. S., & BOWER, G. H. (1982). Remembering script-based text. *Poetics*, *11*, 1-23.
- EINSTEIN, G. O., MCDANIEL, M. A., BOWERS, C. A., & STEVENS, D. T. (1984). Memory for prose: The influence of relational and proposition-specific processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 133-143.
- GRAESSER, A. C., GORDON, S. E., & SAWYER, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. *Journal of Verbal Learning & Verbal Behavior*, *18*, 319-332.
- GRAESSER, A. C., & NAKAMURA, G. V. (1982). The impact of a schema on comprehension and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 53-109). New York: Academic Press.
- GRAESSER, A. C., WOLL, S. B., KOWALSKI, D. J., & SMITH, D. A. (1980). Memory for typical and atypical actions in scripted activities. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 503-515.
- HASTIE, R. (1980). Memory for behavioral information that confirms or contradicts a personality impression. In R. Hastie, T. M. Ostrom, E. B. Ebbesen, & R. S. Wyer (Eds.), *Person memory: The cognitive basis of social perception* (pp. 155-177). Hillsdale, NJ: Erlbaum.
- HEMSLEY, G. D., & MARMUREK, H. H. C. (1982). Person memory: The processing of consistent and inconsistent person information. *Personality & Social Psychology Bulletin*, *8*, 433-438.
- HUNT, R. R., AUSLEY, J. A., & SCHULTZ, E. E., JR. (1986). Shared and item-specific information in memory for event descriptions. *Memory & Cognition*, *14*, 49-54.
- HUNT, R. R., & EINSTEIN, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior*, *20*, 497-514.
- JOHNSON, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning & Verbal Behavior*, *9*, 12-20.
- KINTSCH, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182.
- KINTSCH, W., KOZMINSKY, E., STREBY, W. J., MCKOON, G., & KEENAN, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning & Verbal Behavior*, *14*, 196-214.
- MANDLER, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. Hillsdale, NJ: Erlbaum.
- MARMUREK, H. H. C. (1984). Disposition-based categories: Structure and salience effects on memory. *Journal of Research in Personality*, *18*, 179-194.
- SMITH, D. A., & GRAESSER, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory & Cognition*, *9*, 550-559.
- SRULL, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning & Memory*, *7*, 440-463.
- SRULL, T. K., LICHTENSTEIN, M., & ROTHBART, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 316-345.
- SRULL, T. K., & WYER, R. S., JR. (1989). Person memory and judgment. *Psychological Review*, *96*, 58-83.
- TAYLOR, S. E., & CROCKER, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1, pp. 89-134). Hillsdale, NJ: Erlbaum.
- WYER, R. S., JR., BODENHAUSEN, G. V., & SRULL, T. K. (1984). The cognitive representation of persons and groups and its effect on recall and recognition memory. *Journal of Experimental Social Psychology*, *20*, 445-469.
- WYER, R. S., JR., & GORDON, S. E. (1984). The cognitive representation of social information. R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 2, pp. 73-150). Hillsdale, NJ: Erlbaum.

#### NOTES

1. A number of different terms have been used to describe various types of relatedness. Graesser et al. (1980) used typicality to describe the relatedness of actions to the script. Srull (1981) used congruence to describe the fit between behaviors and expectancies. The term *relevance* will be used consistently to describe the relatedness of actions to a script, even when original authors used a different term. The term *detail expectancy* will be used to describe whether or not the details in actions are the usual, frequently given ones, or whether they are more unusual, nongenerated details. The term *detail expectancy* is not meant to imply that subjects actively instantiate empty argument slots in script actions. *Expectancy* is used, because it was defined by subjects' ratings of details on an expectancy scale.

2. Eighteen subjects also participated in a 1-week delay condition. However, the average proportion of actions recalled was .067, and the average proportion of details recalled was .025. Because floor effects are highly likely, the delay data are not reported.

3. An alternate way to analyze the data would be to conduct a high-versus medium-relevance comparison and a medium-versus low-relevance comparison. Medium-relevance actions were recalled better than high-relevance actions [ $F(1,17) = 4.91$ ,  $MS_e = .02$ ], and medium-relevance actions were also recalled better than low-relevance actions [ $F(1,17) = 5.03$ ,  $MS_e = .01$ ]. Thus, whether the recall of high- and low-relevance actions is compared with the recall of medium-relevance actions in combination or separately, these results support the inverted U hypothesis.

4. The confidence judgments for both hits and correct rejections generally followed the same pattern as did the recognition data. The more closely an action overlapped with the schematic script, the lower the subject's confidence that the action was old or new. Generally, responses to low-relevance actions were given more confidence than responses to high- and medium-relevance actions and unexpected actions were given higher confidence than were expected and generic actions.