

## Congruity Effects in the Survival Processing Paradigm

James S. Nairne

Josefa N. S. Pandeirada

Purdue University

University of Aveiro

Address all correspondence to:

James S. Nairne

Department of Psychological Sciences

Purdue University

703 Third Street

West Lafayette, IN 47907-2081

E-mail: [nairne@psych.purdue.edu](mailto:nairne@psych.purdue.edu)

Running Head: Congruity and Survival Processing

## Abstract

Five experiments were conducted to investigate a proposal by Butler, Kang, and Roediger (2009) that congruity (or fit) between target items and processing tasks might contribute, at least partly, to the mnemonic advantages typically produced by survival processing. In their research, no significant survival advantages were found when words were preselected to be highly congruent or incongruent with a survival and control (robbery) scenario. Experiments 1a and 1b of the present report show that survival advantages, in fact, generalize across a wide set of selected target words; each participant received a unique set of words, sampled without replacement from a large pool, yet significant survival advantages remained. In Experiment 2 we found a significant survival advantage using words that had been preselected by Butler et al. (2009) to be highly unrelated (or irrelevant) to both the survival and control scenarios. Experiment 3 showed a significant survival advantage using word sets that had been preselected to be highly congruent with both scenarios. Finally, Experiment 4 mixed congruent and incongruent words in the same list, more closely replicating the design used by Butler et al. (2009), and a highly reliable main effect of survival processing was still obtained (although the survival advantage for the congruent words did not reach conventional levels of statistical significance). Our results suggest that the null effects of survival processing obtained by Butler et al. (2009) may not generalize beyond their particular experimental design.

Key Words: memory, congruity, survival, evolution

Selection pressures in ancestral environments shaped our capacity to remember (Nairne, 2010; Nairne & Pandeirada, 2008; Sherry & Schacter, 1987). Memory evolved, via the process of natural selection, and consequently was subject to nature's criterion—the enhancement of reproductive fitness. Our laboratory has recently been investigating whether vestiges, or footprints, of ancestral selection pressures remain apparent in the operating characteristics of current memory systems. For example, our laboratory has shown that processing information in terms of its relevance to a survival scenario produces excellent retention—better, in fact, than most well-known encoding procedures (e.g., Nairne, Pandeirada, & Thompson, 2008).

In the typical survival processing procedure, participants are asked to imagine themselves stranded in the grasslands of a foreign land without survival materials. Over the next few months, they will need to find steady supplies of food and water and protect themselves from predators. An unrelated list of words is then presented and the participant's task is to rate the relevance of each word to the imagined survival scenario (Nairne, Thompson, & Pandeirada, 2007). The rating task is followed by a surprise retention test, usually free recall, and performance after survival processing is compared to a variety of control conditions—e.g., processing items for pleasantness, forming a visual image, or in terms of other scenario-based activities (moving to a foreign land, vacationing in the grasslands, etc.). In every case survival processing has produced the best retention (see also Kang, McDermott, & Cohen, 2008; Nairne et al., 2008; Weinstein, Bugg, & Roediger, 2008).

Although survival processing advantages are extremely robust, the proximate mechanisms that produce the retention advantages remain unknown. We have claimed that our memory systems are “tuned” to nature's criterion—remembering information processed for its fitness consequences—but the adaptive problem could presumably be solved in a variety of

ways. For example, processing information in terms of its survival relevance might activate arousal systems in the brain (McGaugh, 2006), lead to distinctive or elaborative processing ( Craik, 2007), or simply induce emotional processing (Phelps, 2006); alternatively, there could be domain-specific modules in the brain that are specifically designed to process and remember certain types of fitness-relevant stimuli and events (Ermer, Cosmides, & Tooby, 2007; Klein, Cosmides, Tooby, & Chance, 2002). However, given the absence of “fossilized” memory traces, and limited information about ancestral environments, building a solid case for cognitive adaptations can be difficult (see Buller, 2005; Nairne, 2010).

The present experiments were designed to investigate a recent proposal by Butler, Kang, and Roediger (2009) that congruity effects between target items and processing tasks might contribute, at least partly, to the oft-replicated survival processing advantage. The congruity effect—or the “principle of congruity”—refers to the general finding that memory performance is enhanced when the encoding context and the to-be-remembered target word form an integrated unit (Schulman, 1974). For example, Craik and Tulving (1975) showed that orienting task questions yielding a “yes” response (e.g., does the target word FRIEND fit the sentence “I meet a \_\_\_\_ in the street?”) produce better retention of the target word than questions producing “no” responses (does the word CLOUD fit the same sentence?). When target words are compatible with the encoding context, a richer and more elaborate representation is produced; in addition, once an integrated unit is formed, the orienting task or encoding question can act as an effective retrieval cue for subsequent recovery of the target (e.g., Moscovitch & Craik, 1976).

The level of congruence between target words and the encoding context can serve as a confounding factor in memory experiments (or as an explanatory mechanism depending on your perspective). Consider the self-reference effect in which people are asked to relate words or trait

adjectives to themselves or to personal autobiographical experiences (Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). Self-referential processing typically produces enhanced retention, compared to other forms of meaningful processing, but one could easily attribute the benefit to a kind of congruity effect. The “self” is a rich, well-practiced, and highly-accessible knowledge structure and the encoding task specifically asks the participant to relate the target word to this structure. Compared to other meaning-based encoding procedures, such as rating an item for pleasantness, self-referential encoding is likely to produce a more highly integrated or congruent encoding between the target and its encoding context. To isolate the “self” as the critical element, one would need to match congruity between the target words and the respective encoding contexts (the “self schema” and pleasantness).

At face value, it is difficult to see how a similar congruity analysis could apply to the survival processing advantage—grasslands scenarios are not particularly familiar or well-practiced, nor is there any direct evidence that they are particularly accessible at retrieval compared to other control scenarios. Moreover, the survival processing advantage remains robust even though there are typically no significant differences in the average relevance ratings given to the target words in the different encoding contexts (e.g., survival versus moving). However, in a pair of experiments Butler et al. (2009) discovered that when the degree of congruity between the target words and the encoding contexts (survival versus a control scenario involving a robbery) is equated, by preselecting target words that were equally matched to the scenarios, the survival processing advantage disappeared. Moreover, words that were selected to be irrelevant, or highly unrelated, to the survival scenario failed to produce survival processing advantages as well.

Participants in the Butler et al. experiments were asked to rate the relevance of word lists to one of two scenarios: the typical grasslands-based survival scenario (Nairne et al., 2007) or to a robbery scenario describing a bank heist. Previous work has shown that survival processing produces better retention than robbery processing when the target words are selected randomly and are nominally unrelated (Kang et al., 2008). Processing scenario was manipulated between-subjects, but each participant received three lists of target words (presented either in a blocked or random fashion). One list contained words that were preselected by independent participants as highly relevant to the survival scenario, one list contained words that were highly relevant to the robbery scenario, and the third list contained words that were highly unrelated to either scenario. A surprise recall test revealed substantial congruity effects—the survival (robbery) words processed via the survival (robbery) scenario were remembered much better than the robbery (survival) words processed via the survival (robbery) scenario, but no overall differences were found between survival and robbery processing (for any word class). These results are important because they constrain the generality of the survival processing advantage.

Following Butler et al. (2009), the current experiments were designed to investigate the role that congruity effects potentially play in the survival processing paradigm. Experiments 1a and 1b attempt to replicate the survival processing advantage under conditions in which each participant receives a completely novel set of target words—this experiment was designed to reduce the possibility that uncontrolled congruity effects might have been present in previous experiments using fixed word lists as target stimuli. Experiments 2 – 4 attempt to replicate the Butler et al. (2009) results, using the same set of scenarios and items, but using somewhat different experimental designs. To anticipate the results, significant survival advantages were

obtained in every experiment, although congruity remained a powerful determinant of performance.

### Experiment 1a and 1b

The majority of studies conducted using the survival processing paradigm have employed unrelated word lists sampled randomly from various norms (e.g., Van Overschelde, Rawson, & Dunlosky, 2004). Typically, a single set of words is used in an experiment such that, across participants, the same words are rotated through the various conditions (e.g., survival and control conditions). Because everyone is tested on identical items, little attention has been given to the characteristics of the words or their potential relevance (or congruity) to the encoding scenarios. Consequently, as Butler et al. (2009) suggested, it is possible that the words used in previous experiments might have contained “a greater number of words that are congruent with the survival processing scenario than other processing tasks” (p. 1484).

Our first two experiments test whether the survival advantage generalizes to a very wide sample of words. To accomplish this end, each participant in Experiments 1a and 1b received a completely novel set of words—no target word was repeated across conditions or participants within an experiment. List items were sampled randomly, without replacement, from the Paivio, Yuille, and Madigan (1968) norms. In Experiment 1a survival processing was compared to a pleasantness rating task; Experiment 1b compared survival processing to the moving scenario used previously by Nairne et al. (2007). Demonstrating a survival processing advantage under these conditions would dramatically reduce the chances that it could be plausibly attributed to an item-selection artifact.

### *Method*

*Participants and apparatus.* Twenty eight Purdue undergraduates participated in each experiment either in exchange for partial credit in an introductory psychology course (Experiment 1a) or for a small monetary compensation (Experiment 1b). Everyone was tested in sessions lasting approximately 30 min. Up to three participants were tested in the same session. Stimuli were presented and controlled by personal computers.

*Materials and design.* Twenty eight lists of words were drawn randomly from the set of 925 nouns reported in Paivio et al. (1968); each list contained 32 different words for a total of 896 to-be tested words across participants. An additional set of words was sampled to be used as practice items. This selection procedure was employed twice to compose the 28 lists used in each experiment.

A within-subject design was used in both experiments. In Experiment 1a, participants were instructed to make either a survival or a pleasantness rating about each of the 32 target words. The rating tasks were distributed evenly across the list with the following constraints: no more than two ratings of a given type could occur consecutively, and the same number of survival and pleasantness decisions were required in each half (i.e., eight survival and eight pleasantness decisions in the first half, and the same for the second half). Task order was counterbalanced across participants to ensure that a survival and a pleasantness decision occurred the same number of times in each position of the list. In Experiment 1b, participants rated the relevance of 16 words to the survival (S) scenario and the relevance of a second set of 16 words to the moving (M) control. Rating condition was blocked in trials of eight words in the form SMSM or MSMS. Half of the lists were randomly assigned to each counterbalancing version of the encoding task and, within each list, words were randomly assigned to each block and condition.

*Procedure.* On arrival in the laboratory, people were assigned to one of the word lists. In Experiment 1a, people were told they would be rating words in two ways—either for pleasantness or with respect to a survival situation. In Experiment 1b, participants received general instructions informing them that they would be required to rate words with respect to particular scenarios. Then, either the survival or moving instructions appeared depending on the counterbalancing condition. The specific instructions for each rating condition were as follows:

*Survival:* For the SURVIVAL situation, please imagine that you are stranded in the grasslands of a foreign land, without any basic survival materials. Over the next few months, you'll need to find steady supplies of food and water and protect yourself from predators. We would like you to rate how relevant the word would be for you in this survival situation. The scale of relevance ranges from one to five, with one (1) indicating totally irrelevant and five (5) signifying extremely relevant. Some of the words may be relevant and others may not - it's up to you to decide.

*Pleasantness:* For the other dimension, we would like you to rate the PLEASANTNESS of the word. The scale of pleasantness ranges from one to five, with one (1) indicating totally unpleasant and five (5) signifying extremely pleasant. Some of the words may be pleasant and others may not - it's up to you to decide.

*Moving:* In this task, we would like you to imagine that you are planning to move to a new home in a foreign land. Over the next few months, you'll need to purchase a new house and find help transporting your belongings. Please rate how relevant each of these words would be for you in this moving situation. Some of the words may be relevant and others may not - it's up to you to decide. Remember, the scale of relevance ranges from one to five, with one indicating totally irrelevant and five signifying extremely relevant.

In Experiment 1a, a short practice session of 6 items preceded the main rating session. Each word was presented with a question that specified the rating decision to be made for that specific word (“How pleasant is this word?”, or “How relevant is this word to the survival situation?”). The rating scale was also presented along with the word. It ranged from one to five and was labeled appropriately (1 = totally irrelevant / totally unpleasant; 5 = extremely relevant / extremely pleasant). In Experiment 1b a short practice session containing three words was included at the beginning of the first and second blocks to ensure that people understood the two rating scenarios. In this experiment the rating scale and corresponding labels were the same for both encoding tasks (1 = totally irrelevant; 5 = extremely relevant). In both experiments, and in all subsequent experiments, people were instructed to try and use the full rating scale.

In both experiments the target words were presented individually for 5 s in the center of the screen; target words always remained on the screen for the full 5 s, irrespective of when a rating response was entered. Participants produced their responses by clicking on the button that corresponded to the rating of their choice. After the last word was rated, instructions appeared for a short distractor task. For this task, seven digits, ranging between zero and nine, were presented sequentially for one s apiece and participants were required to recall the digits in order by typing responses into a text box. The digit recall task proceeded for approximately two min. Recall instructions then appeared. Participants were instructed to write down the earlier-rated words, in any order, on a response sheet. The final recall phase proceeded for 5 min and participants were asked to draw a line on the recall sheet, under the last recalled word, after each min of recall. A clock was displayed on the computer monitor and a “beep” sounded every min signaling the participants to draw the line. Using this procedure allows one to construct cumulative recall curves, but they are not reported here.

### *Results and Discussion*

The level of statistical significance, unless otherwise noted, was set at  $p < .05$  for all comparisons. In both experiments participants had little difficulty producing relevance ratings for the individual stimuli within the allotted time, and no significant differences in completion rates were found between conditions. Because of the small number of unrated trials and to avoid item selection problems, we left the retention data described below unconditionalized.

The data of main interest are presented in Figure 1. The left hand side of the figure shows proportion correct recall for words rated initially for survival and pleasantness (Experiment 1a) and the right hand side shows memory performance for survival and moving (Experiment 1b). For the data in Experiment 1a, a repeated measures ANOVA confirmed a survival advantage over pleasantness [ $F(1, 27) = 7.36$ ,  $MSe = .009$ ,  $\eta^2_p = .21$ ]. Out of the 28 participants, 20 recalled more words rated for survival than for pleasantness, 7 recalled more pleasantness words, and there was 1 tied score. A similar pattern was found in Experiment 1b: People remembered significantly more words in the survival condition than in the moving condition [ $F(1, 27) = 4.72$ ,  $MSe = .008$ ,  $\eta^2_p = .15$ ]. For the 28 participants, 16 recalled more survival words, 8 remembered more moving words and there were 4 ties.

Average rating and response time data are shown in Table 1 for each condition in each experiment. In Experiment 1a, pleasantness ratings were significantly higher than the relevance ratings produced in the survival condition [ $F(1, 27) = 4.79$ ,  $MSe = .120$ ,  $\eta^2_p = .151$ ]. In Experiment 1b, the survival ratings did not differ significantly from those given to the moving scenario [ $F(1, 27) = 1.33$ ,  $MSe = .091$ ,  $\eta^2_p = .05$ ]. For response times, participants were slower at producing a survival rating than a pleasantness rating [ $F(1, 27) = 49.29$ ,  $MSe = 43350.30$ ,  $\eta^2_p = .65$ ], but no significant differences in response time were found between survival and moving [ $F(1, 27) =$

1.11,  $MSe = 76588.5$ ,  $\eta^2_p = .04$ ]. As in previous survival processing experiments, neither the rating nor the response time data seem capable of explaining the significant survival processing advantages found in free recall across the experiments.

Together, these two experiments show that survival processing can produce mnemonic advantages compared to a standard form of deep processing (pleasantness rating) and a non-fitness relevant schematic encoding task (a moving scenario) even when target items are sampled from a large word pool. Again, each participant in each experiment received a different set of words; consequently, the chances that prior demonstrations of the survival advantage might have been artifactual, a consequence of some kind of item-selection problem, are reduced. Of course, these data do not rule out congruity accounts—one can simply assume that all words, on average, are more congruent with survival processing than with relevant controls. We consider this account unlikely, as noted above, because the typical grasslands scenario is neither familiar nor well-practiced. Our remaining experiments examine survival processing under conditions in which congruity (or its absence) was controlled directly, through matched selection of target words, following up on recent research by Butler et al. (2009).

## Experiment 2

One of the striking findings of Butler et al. (2009) was their failure to find a survival processing advantage for words deemed irrelevant to either the survival or the robbery processing scenario. Most survival processing experiments have used unrelated word lists, selected randomly from norms, and survival advantages have generally been found for words irrespective of their assigned relevance rating (e.g., Butler et al., 2009, experiment 1; Nairne et al., 2007). Even words given the lowest relevance rating (1 on a 5-point scale) often show a survival processing advantage when ratings are matched across conditions. Experiment 2 was

designed to replicate the Butler et al. (2009) findings using a somewhat different experimental design.

As described earlier, each participant in Butler et al. (2009) received three lists of target words (presented either in a blocked or random fashion). One list contained words that were preselected to be highly relevant to a survival scenario, one list contained words that were highly relevant to a robbery scenario, and the third list contained words that were irrelevant to either scenario. Thus, for a given participant, who received only the survival or robbery scenario, two-thirds of the words were irrelevant to the assigned processing scenario (averaging lower than 1.5 on the 5-point scale) and one-third were highly congruent words (averaging higher than 4.0 on the 5-point scale). In our Experiment 2 participants received only words that were irrelevant to their assigned scenario. Survival versus robbery processing was manipulated between-subjects. The question of main interest asked whether survival processing would continue to produce recall advantages under conditions in which only words deemed irrelevant to the processing scenarios are used.

### *Method*

*Participants and apparatus.* One hundred and ten Purdue students participated in this experiment in exchange for partial credit in an introductory psychology course. Participants were tested individually or in groups of up to 4 people in sessions lasting approximately 30 min. Stimuli were presented and controlled by personal computers.

*Materials and design.* All stimulus materials were taken directly from Butler et al. (2009; see pages 1485 and 1486). A mixed design was used with processing scenario (i.e., survival versus robbery) manipulated between-subjects and list (i.e., type of irrelevant list) as a within-subject variable. All participants were asked to rate the relevance of 30 words to either a survival

or a robbery scenario (n = 55 in each group). Fifteen of the to-be-rated words were irrelevant to both the survival and the robbery scenarios; the other fifteen words were either survival- or robbery-relevant, but irrelevant to the assigned processing condition. Thus, participants in the survival condition received the Butler et al. robbery list and participants in the robbery condition received their survival list. All of the words were therefore “irrelevant” to the assigned rating task, but half of the words were related in the sense that they shared relevance to an unseen scenario. The rating task was followed immediately by a short digit-recall task prior to the final unexpected free-recall task. Except for the rating scenario and half of the to-be rated words, all aspects of the design, including timing, were held constant across participants.

Within each list words were randomly intermixed with the following constraints: 1) words from the irrelevant list (i.e., the words common to both conditions) were presented in the exact same position of the list in both conditions, and 2) the same number of these irrelevant words was presented in each half of the list. Five irrelevant words, also provided by Butler et al., were used as practice words. All lists of words were equated for frequency, length and imageability by Butler et al. (2009).

*Procedure:* Participants were assigned to one of two conditions based on their arrival time at the laboratory. As in Butler et al., the instructions used in each condition were as follows:

*Survival:* In this task we would like you to imagine that you are stranded in the grasslands of a foreign land, without any basic survival materials. Over the next few months, you'll need to find steady supplies of food and water and protect yourself from predators. We are going to show you a list of words, and we would like you to rate how relevant each of these words would be for you in this survival situation. Some of the words may be relevant and others may not – it's up to you to decide.

*Robbery:* In this task we would like you to imagine that you are leading a heist of a well-guarded bank. Over the next few months, you'll need to find people to help you, make a plan, and gather any supplies you might need. We are going to show you a list of words, and we would like you to rate how relevant each of these words would be for you in this robbery situation. Some of the words may be relevant and others may not – it's up to you to decide.

This experiment replicated the procedural details of Experiments 1a and 1b including form and timing of presentation of the stimuli. Participants rated the words using the same 5-point scale in both conditions, with 1 indicating totally irrelevant and 5 signifying extremely relevant. After rating the words, participants performed the distractor digit memory task for about 2 min and were then surprised with the free recall task. In this experiment participants were given a 10 min recall period; they were instructed to write down the earlier rated words, in any order, on a response sheet.

### *Results and Discussion*

As in the previous experiments, participants successfully rated over 99% of the words within the allotted time, and no differences were found between conditions.

The data of main interest are shown in Figure 2, which presents proportion correct recall for each type of list (common versus scenario-dependent) in the two rating conditions. A mixed ANOVA with condition (i.e., survival vs. robbery) as a between-subject variable and list type as a within-subject variable revealed a significant main of condition [ $F(1, 108) = 11.68$ ,  $MSe = .025$ ,  $\eta_p^2 = .098$ ], a main effect of list [ $F(1, 108) = 18.98$ ,  $MSe = .017$ ,  $\eta_p^2 = .149$ ], and no reliable condition X list interaction ( $F < 1$ ). The significant survival processing advantage seen here contrasts with Butler et al. (2009) who obtained no survival advantage for any list type in their experiments. Note that we used exactly the same word sets as those authors, except that in the

present case participants received only the “irrelevant” words. Our results suggest that the null effect of survival processing obtained by Butler et al. (2009) may have been due to the mixing of highly congruent and incongruent words in the same list. We return to this issue in Experiment 4.

The other finding of interest was the recall advantage for the scenario-dependent words over the common irrelevant words. This advantage is somewhat difficult to interpret, however, because one is comparing across different word sets (e.g., survival-relevant and common irrelevant). Butler et al. (2009) were careful to equate their word lists on most mnemonic dimensions (e.g., frequency and imageability), but item-selection artifacts are difficult to rule out completely. In addition, words from the scenario-dependent lists share commonalities, by virtue of their relevance to the unassigned scenario, that might contribute to the recall advantage. One can think of these word sets as comprising *ad hoc categories*—e.g., things that are relevant in planning a bank heist—and participants may have been sensitive to relationships among the words that ultimately benefitted recall.

Rating and response time data for each type of list in each condition are presented in Table 2. A mixed ANOVA on the rating data revealed significant main effects for list and condition [ $F(1, 108) = 56.36$ ,  $MSe = .077$ ,  $\eta_p^2 = .343$ ;  $F(1, 108) = 20.11$ ,  $MSe = .448$ ,  $\eta_p^2 = .157$ , respectively], but also a reliable condition X list interaction [ $F(1, 108) = 22.29$ ,  $MSe = .077$ ,  $\eta_p^2 = .171$ ]. Specifically, words from the scenario-dependent lists (the survival and robbery words) were rated as more relevant than the common words, but the difference was larger in the robbery processing condition. In addition, words were rated as more relevant to the robbery scenario than to the survival condition for both the common and the scenario-dependent words, as confirmed by additional tests [ $F(1, 108) = 5.97$ ,  $MSe = .240$ ,  $\eta_p^2 = .052$ , and  $F(1, 108) = 32.62$ ,  $MSe = .285$ ,  $\eta_p^2 = .232$ , respectively]. Note that this last result is the opposite of what one would expect from

a congruity analysis—words were rated as more relevant to the robbery scenario than to the survival scenario, but survival processing produced better retention.

Turning to the response time data, the mixed ANOVA showed a marginally reliable main effect of list [ $F(1, 108) = 3.66$ ,  $MSe = 43222.11$ ,  $\eta_p^2 = .033$ ], reflecting a slower rating time for the common list than for the scenario-dependent lists. Neither the main effect of scenario nor the condition X list interaction was significant [ $F < 1$ , and  $F(1, 108) = 2.05$ ,  $MSe = 43222.11$ ,  $\eta_p^2 = .019$ , respectively].

In Experiment 2, again, all of the target words rated by a given participant were ostensibly unrelated (or irrelevant) to their assigned processing scenario. These words were selected by Butler et al. (2009), based on independent ratings, in an effort to eliminate any uncontrolled congruity between the target words and their assigned processing scenarios. When words are selected randomly, as in previous research, certain words, by chance, might be more related to the survival scenario than to the various controls. The results of the current experiment, in combination with the findings of Experiments 1a and 1b, substantially lower the chances that previous demonstrations of the survival processing advantage were due to such uncontrolled congruity effects.

### Experiment 3

Butler et al. (2009) also found a null effect of survival processing for highly-congruent words—that is, words that had been rated by an independent group of participants as highly related to either the survival or the robbery processing scenarios. If previous demonstrations of the survival advantage were affected by uncontrolled congruity—i.e., random word lists contained, over average, more words that were congruent with survival than robbery processing—then using only congruent words might be expected to reduce or eliminate the effect

as well. This is exactly what Butler et al. (2009) found: Survival processing of “survival words” produced no significant recall advantages over robbery processing of “robbery words.”

Experiment 3 attempts to replicate the Butler et al. (2009) findings, using the same scenarios and word sets, but in a design using only congruent words.

### *Method*

*Participants and apparatus.* One hundred and fourteen Purdue undergraduates participated for credit in an introductory psychology course. Stimuli were presented and controlled by personal computers in sessions lasting approximately 30 min. Sessions were conducted in groups of up to 4 participants.

*Materials and design.* The survival and robbery lists were used in this experiment. A simple within-subject design was employed: Each participant rated 14 words from the survival list for the survival scenario (S), and 14 words from the robbery list for the robbery scenario (one word from each list was randomly eliminated in order to get two blocks with equal number of words from each list). Within each list words were divided in two blocks of seven words apiece, and order of presentation within each block was randomly determined and the same for all participants. The rating task was blocked in trials of seven words in the form SRSR or RSRS. Word block presentation was also counterbalanced. As in the previous experiments word presentation was followed by a short digit recall task and then participants were surprised with the free recall test.

*Procedure.* The procedure used in this experiment was the same as in Experiment 1b except for the control scenario and the duration of the recall task. In this experiment, the robbery scenario from the previous experiment was used as the control for survival processing and

participants were allowed 10 min for the free recall task. All other aspects of the procedure were as described for Experiment 1b.

### *Results and Discussion*

Once again, participants had no trouble completing the ratings within the 5-s presentation window. The data of main interest are shown in Figure 3, which shows proportion correct recall for the words rated with respect to the survival or robbery scenarios. A repeated measures ANOVA confirmed that survival processing produced significantly better retention than robbery processing ( $F(1, 113) = 7.13$ ,  $MSe = .015$ ;  $\eta_p^2 = .059$ ).

Table 3 shows the rating and response time data for each condition. As expected, given that the words were preselected to be highly congruent with their respective scenarios, the average ratings were quite high. Words were given slightly higher ratings in the survival condition than in the robbery scenario (averages of 4.03 and 3.92); this difference was statistically reliable ( $F(1, 113) = 6.94$ ,  $MSe = .091$ ;  $\eta_p^2 = .058$ ). With respect to congruity, higher ratings represent a better fit to the scenario and could possibly account for the better memory performance in the survival condition. In the Butler et al. (2009) study, no significant ratings differences were found between the survival- and robbery-congruent words. So, additional analyses were conducted to try to rule out rating as a determinant of the survival recall advantage found here. First the difference in average rating between survival and robbery was calculated for each participant. Participants whose rating difference between the two conditions was smaller than .15 were selected; this process resulted in a set of 32 participants with an average rating of 4.0 in both conditions ( $t < 1$ ). (Each counterbalancing version was represented equally in the sample.) For these participants, proportion of recall for survival was .71 and .63 for robbery, a reliable difference ( $t(31) = 2.36$ ;  $p < .03$ ). It is worth noting as well that in Experiment 2 words

were rated as more relevant to the robbery scenario than to the survival scenario, but survival processing produced better retention. No significant differences were found in response time.

These results clearly demonstrate that it is possible to obtain significant survival processing advantages using congruent stimuli. Both the survival and robbery words were preselected to match congruity between scenario conditions yet, in contrast to the findings of Butler et al. (2009), a significant survival processing advantage was still obtained. In combination with the results of Experiment 2, our data suggest that the null effects of survival processing found by Butler et al. (2009) may be specific to their particular experimental design, one in which participants received both highly incongruent and congruent words in the same experimental session.

#### Experiment 4

As noted throughout, in their investigation of congruity effects in the survival processing paradigm Butler et al. (2009) used a mixed design in which both congruent and incongruent words were processed with respect to either a survival or robbery scenario. Substantial congruity effects were found—congruent words were remembered much better than incongruent words—but no survival advantages were detected for either the congruent or the incongruent (irrelevant) conditions. In our experiments, which used exactly the same stimuli and scenarios as Butler et al. (2009), survival advantages were found for both congruent and incongruent words. Of course, participants received only one of the item types in a given experiment—incongruent (or irrelevant) words (Experiment 2) and congruent words (Experiment 3). In Experiment 4, participants were asked to make either survival or robbery decisions about both congruent and incongruent items in the same session.

#### *Method*

*Participants and apparatus.* One hundred and seventy two Purdue undergraduates took part in the experiment in exchange for partial credit in an introductory psychology course. Eighty six were assigned to each condition. Participants were tested in groups of up to 4 people. Stimuli were presented and controlled by personal computers.

*Design and Materials.* A mixed design similar to Experiment 2 was used in this experiment: Scenario (i.e., survival versus robbery) was manipulated as a between-subject variable and item type (i.e., congruent and incongruent) as a within-subject variable. All participants were asked to rate the relevance of 30 words to either the survival or robbery scenario. Half of the target words were incongruent to both scenarios (the common “irrelevant” words used in Experiment 2) and the remaining words were the congruent words used in Experiment 3. Participants receiving the survival scenario were given the survival-relevant words and the robbery group received the robbery-relevant words. The congruent and incongruent (or irrelevant) words were mixed together throughout the experimental session in a manner described in Experiment 2. The rating task was followed immediately by a short digit-recall task prior to the final unexpected free-recall task. Except for the rating scenario, and the congruent word pools, all other aspects of the design, including timing, were held constant across participants.

*Procedure:* Participants were randomly assigned to one of two conditions on their arrival at the laboratory. The instructions used in each condition were the same as in the previous experiment. Again, all other aspects of the procedure were the same as in Experiment 2.

### *Results and Discussion*

Ratings were provided for over 99% of the presented words, and the number of unrated words did not differ significantly between groups.

Figure 4 shows proportion correct recall for each type of list in each condition. A mixed ANOVA with condition (i.e., survival vs. robbery) as a between-subject variable and list type (irrelevant list vs. congruent list) as a within-subject variable revealed a significant main effect of condition [ $F(1, 170) = 10.59$ ,  $MSe = .024$ ,  $\eta_p^2 = .059$ ], a main effect of list [ $F(1, 170) = 261.39$ ,  $MSe = .014$ ,  $\eta_p^2 = .61$ ], and a marginally reliable condition X list interaction [ $F(1, 170) = 3.55$ ,  $MSe = .014$ ,  $\eta_p^2 = .02$ ]. Survival processing produced better recall performance than robbery processing, and substantially more congruent words were recalled than incongruent (irrelevant) words. However, there was a marginally reliable interaction suggesting that survival processing may have had a larger effect for the irrelevant items than for the congruent items. Indeed, post-hoc analyses comparing recall as a function of word type revealed a highly significant survival advantage for the incongruent words [ $F(1, 170) = 14.24$ ,  $MSe = .019$ ,  $\eta_p^2 = .077$ ], but the effect failed to reach significance for the congruent words [ $F(1, 170) = 2.03$ ,  $MSe = .019$ ,  $\eta_p^2 = .012$  ( $p < .08$  one-tailed)].

Rating and response time data for each type of word in each condition are presented in Table 4. The mixed ANOVA on the rating data revealed a significant main effect of item type [ $F(1, 170) = 3916.28$ ,  $MSe = .123$ ,  $\eta_p^2 = .958$ ], with the congruent words rated higher than the irrelevant words. The condition X list interaction was also reliable [ $F(1, 170) = 5.72$ ,  $MSe = .123$ ,  $\eta_p^2 = .033$ ], reflecting a larger rating difference between word type for the robbery condition than for the survival condition. The main effect of condition was not reliable ( $F < 1$ ). For the response time data, the mixed ANOVA revealed only a reliable main effect of item type [ $F(1, 170) = 41.28$ ,  $MSe = 74085.72$ ,  $\eta_p^2 = .195$ ], with participants responding slower to the irrelevant words. Neither the effect of condition nor the list X condition interaction reached

significance [ $F(1, 170) = 1.67$ ,  $MSe = 233130.29$ ,  $\eta_p^2 = .010$ , and  $F(1, 170) = 3.32$ ,  $MSe = 74085.72$ ,  $\eta_p^2 = .019$ ].

The results of Experiment 4 replicated those of Experiments 2 and 3 in demonstrating strong survival processing advantages, using the same word sets as Butler et al. (2009), although the survival advantage for the congruent words failed to reach significance in a post-hoc test. Experiment 4 differed from Experiment 3, in which a reliable survival advantage was found for congruent words, in that people made rating decisions about both congruent and irrelevant words in the same session. In the Butler et al. (2009) studies, congruent and irrelevant words were mixed in the same session as well (either blocked or intermixed) and no survival advantages were found for any word set. The present data suggest that mixing the word types together might reduce (and perhaps eliminate) the survival advantage for congruent words, but we still found strong survival advantages for the irrelevant words.

What explains the discrepancy? The major difference between Butler et al. (2009) and the current Experiment 4 was the proportion of congruent to irrelevant words in the session. Whereas we used only the irrelevant and the congruent word sets from Butler et al., which produced equal numbers of each word type, participants in the Butler et al. studies received one set of congruent words and *two* sets of irrelevant words (the irrelevant list and the scenario-dependent but irrelevant list). It is unclear why the proportion of word types within the list should matter, although rating a greater number of incongruent words may have enhanced the perceived fit between the congruent words and their processing scenario. The relevance ratings for the congruent words were somewhat higher in Experiment 4 than in Experiment 3 and the response times were somewhat lower as well. However, it remains a mystery why Butler et al. (2009) were unable to find survival advantages for any word class. The average ratings for the

incongruent words were lower in Experiment 4 than in Experiment 2, perhaps because of the inclusion of congruent words in the same list, but strong survival processing advantages were still found for this word set.

The other major finding of Experiment 4, replicating Butler et al. (2009), was the strong overall effect of congruity. Survival processing of the congruent survival words produced much better recall than survival processing of the irrelevant words; the same congruency pattern was found for robbery processing. As both Butler et al. (2009) and our experiments demonstrate, it is not the word set *per se* that produces this retention pattern, but the fit (or congruency) between the word set and its respective processing scenario. Obviously, congruent processing is an extremely powerful determinant of later retention.

#### General Discussion

Previous work using the survival processing paradigm has established that processing information in terms of its survival value enhances retention relative to standard “deep processing” tasks (e.g., Nairne et al., 2008; Nairne et al., 2007) and other matched control scenarios. For example, Nairne, Pandeirada, Gregory and Van Arsdall (2009) used a hunting scenario in which people were asked to rate the relevance of words to hunting big game, trapping small animals, or fishing in a nearby lake. In the fitness-relevant version of the scenario, these activities were introduced as a means to secure food for survival; in the control scenario, the same activities were described as necessary to win a “hunting contest.” Participants in the fitness-relevant condition showed the best retention, even though everyone was rating the relevance of words to essentially the same activities.

Although our experiments have been motivated by an evolutionary perspective, we have been cautious about interpreting the retention advantages. We can be relatively certain that our

memory systems evolved because of their fitness-enhancing properties, such as remembering the actions and locations of predators, but nature could have solved the problem of remembering fitness-relevant information in a variety of ways. For example, fitness-relevant processing might increase arousal, induce emotional processing, produce a richer and more elaborate memory trace, or some combination of these effects. At the same time, our memory systems might be uniquely “tuned” for survival processing, perhaps through the activation of special adaptations that control retention in fitness-relevant situations. Enhanced retention of information processed for its fitness consequences is an *a priori* prediction of an evolutionary analysis, but it is exceedingly difficult to build a strong case for cognitive adaptations (see Andrews, Gangestad, & Matthews, 2002; Nairne, 2010).

In the present case, we were interested in exploring the role of congruity between materials and processing tasks in the survival processing paradigm. Our experiments were conducted in response to recent research by Butler et al. (2009), who reported that the survival processing advantage disappeared when congruity was controlled in a mixed list design. As noted throughout, participants in their experiments rated the relevance of target words that were preselected (based on independent ratings) to be congruent or incongruent with either a survival scenario or a robbery-based control scenario. Scenario was manipulated between-subjects, but all participants received both the congruent and incongruent words which were rated in either a blocked (Experiment 2) or an intermixed (Experiment 3) fashion. Neither of their experiments produced any survival processing advantages, although large effects of congruity were found.

Our results suggest that the null effects of survival processing obtained by Butler et al. (2009) may not generalize beyond their particular experimental design. Experiment 1 showed that the basic survival processing advantage does not depend on the selected target words; each

participant received a unique set of words, sampled without replacement from a large pool, yet significant survival processing advantages remained. In Experiment 2 we found a significant survival advantage for words that had been preselected by Butler et al. (2009) to be highly unrelated (or irrelevant) to the survival scenario. Experiment 3 also showed a significant survival advantage for words sets that had been preselected to be highly congruent with the survival scenario. Finally, Experiment 4 mixed congruent and incongruent words in the same list, more closely replicating the design used by Butler et al. (2009), and a highly reliable main effect of survival processing was still obtained (although the survival advantage for the congruent words did not reach conventional levels of statistical significance).

It is unclear why we were unable to replicate most of the null effects reported by Butler et al. (2009), although it is easy to see how the nature of the experimental design might matter. When only one word type is employed in a list—e.g., only congruent words—people are likely to discriminate more finely among the words during the rating task—that is, to find subtle connections between the to-be-rated word and the processing scenario—compared to when multiple word classes are included. The task instructions encouraged participants to use the full rating scale, so subtle differences among highly-congruent words likely became more salient when only congruent words were available for processing. Overall, when the lists were mixed people treated the incongruent words as less related to survival (or robbery), and the congruent words as more related to the scenario, than they did otherwise. Ratings for the irrelevant (common) words in Experiment 2 averaged 2.22 on a five-point scale whereas ratings for those same words in Experiment 4, which included congruent words, averaged 1.94. Similarly, the ratings for congruent words in Experiment 3 averaged 3.98 and 4.31 in the mixed list design of Experiment 4. People were also slower to make their ratings when only congruent or irrelevant

words were included in a list. For example, response latency averaged 2365 ms for congruent words in Experiment 3, but latency for those same words decreased to 2231 ms in Experiment 4. Perhaps when it is easy to assign relevance to a particular word—clearly relevant or irrelevant to the scenario—the amount of scenario-based processing decreases.

It is worth noting that the ratings obtained in Experiment 4 were very similar to the ratings reported by Butler et al. (2009). For the congruent words, Butler et al. reported an average rating of 4.35 (compared to our 4.31) and for the irrelevant words their average rating was 1.85 (compared to our 1.94). It seems unlikely that these small rating differences could be responsible for the contrasting overall data patterns (our strong main effect of survival processing) but it remains a possibility. Again, Butler et al. (2009) used word lists with more incongruent than congruent words. Of course, mixed lists can also change retrieval strategies, perhaps leading to more or less reliance on the original encoding scenario as a retrieval cue during recall. Recall by word class can lead to differential output interference between word types as well, although Butler et al. (2009) were able to reject output order as an explanation for their null results. At this point we are unable to offer any definitive explanation for the discrepancies between our findings and those of Butler et al. (2009), although clearly one important answer lies in the nature of the experimental design.

One striking finding common to both investigations is the effect that congruity has on overall recall. In most cases, processing words that were congruent with a processing scenario greatly enhanced their later recall—in fact, the congruity effect seen in Experiment 4 was significantly larger than the size of the survival processing advantage. Of course, the mnemonic power of congruity has been demonstrated repeatedly over the years (e.g., Craik & Tulving, 1975), although its mnemonic locus is not particularly well understood. The standard

interpretation is that a good fit between the processing task and the target word affords a richer and more elaborate encoding, along with a beneficial retrieval plan. For example, suppose that the orienting task required people to decide whether target words fit the category “furniture.” If some of the target words were pieces of furniture, then people could use the category structure at retrieval to generate possible recall candidates—obviously, this would benefit words that fit the category more than words that did not. In the present case, participants might have used the different scenarios as retrieval cues, thus enhancing the chances that congruent target words would come to mind.

With respect to survival processing in general, the current results reduce the chances that uncontrolled congruity explains the survival processing advantages seen in previous research. As reviewed above, reliable survival processing advantages were found for all word sets across our experiments—words that were preselected to be congruent or incongruent to the survival scenario. Previous analyses of the relationship between recall and relevance rating reinforce this conclusion (e.g., Nairne et al., 2007). There is often a main effect of rating on recall—words given higher ratings tend to be recalled better—but no interaction between the survival and control conditions (e.g., Butler et al., 2009). The size of the survival advantage does not seem to depend importantly on the ratings given to the individual words. In the present experiments, in which the level of congruity was determined independently (see Butler et al., 2009), the survival advantage was reduced somewhat for the congruent words, although this trend was most noticeable only when congruent and incongruent words were intermixed in the same list. High levels of congruity could mask the advantages of survival processing, perhaps because people rely heavily on congruity as a basis for a retrieval strategy, but it is clearly possible to

demonstrate significant survival advantages when the fit between the words and the scenarios is high (e.g., Experiment 3).

From an evolutionary perspective, it might seem strange that all stimuli bathed in the spotlight of survival processing seem to receive some kind of mnemonic boost. One might expect survival processing only to benefit items considered survival-relevant—after all, what adaptive edge is gained by remembering items that are deemed irrelevant to fitness? As we have discussed elsewhere (Nairne, 2010; Nairne & Pandeirada, 2010), in natural settings fitness-relevant stimuli will typically receive the spotlight of processing attention. Irrelevant events, unlike in the laboratory, will either be ignored or processed with less vigor. The survival processing paradigm forces participants to consider the fitness-relevant properties of all presented stimuli, regardless of whether they are ultimately found relevant or not. Moreover, the fitness properties of a stimulus are likely to be context-dependent. As Nairne and Pandeirada (2008) put it: “food is survival relevant, but more so at the beginning of a meal than at its completion; a fur coat has high s-value at the North Pole, but low at the Equator” (p. 240). Even what seems to be a completely irrelevant stimulus, such as a pencil, can become fitness-relevant under the right circumstances (e.g., a pencil can be used as a weapon in an attack). For this reason we have suggested that survival *processing* is the key to long-term enhancement, although stimuli that are naturally fitness-relevant might show better retention as well (see Nairne, 2010, for further discussion).

We believe that empirical investigations of memory benefit from adopting a functional/evolutionary perspective. To understand the operating characteristics of a cognitive system, it is useful to consider the selection pressures that led to its development. Like any other biological system, our capacity to remember evolved ultimately because it satisfied nature’s

criterion—the enhancement of reproductive fitness. The survival processing advantages seen in the present experiments are consistent with such an evolved system, regardless of the proximate mechanisms that control retention.

## References

- Andrews, P. W., Gangestad, S. W., & Matthews, D. (2002). Adaptationism – how to carry out an exaptationist program. *Behavioral and Brain Sciences*, *25*, 489-504.
- Buller, D. J. (2005). *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. Cambridge, MA: MIT Press.
- Butler, A. C., Kang, S. H. K., & Roediger, H. L. (2009). Congruity effects between materials and processing tasks in the survival processing paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1477-1486.
- Craik, F. I. M. (2007). Encoding: A cognitive perspective. In H. L. Roediger, Y. Dudai & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 129-135). New York, NY: Oxford University Press.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268-294.
- Ermer, E., Cosmides, L., & Tooby, J. (2007). Functional specialization and the adaptationist program. In S. W. Gangestad & J. A. Simpson (Eds.), *The evolution of mind: Fundamental questions and controversies* (pp. 86-94). New York, NY: Guilford Press.
- Kang, S. H. K., McDermott, K. B., & Cohen, S. M. (2008). The mnemonic advantage of processing fitness-relevant information. *Memory & Cognition*, *36*, 1151-1156.
- Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychological Review*, *109*, 306-329.

- Masson, M. E. J., & Loftus, G. R. L. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203-220.
- McGaugh, J. L. (2006). Make mild moments memorable: Add a little arousal. *Trends in Cognitive Sciences*, *10*, 345-347.
- Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 447-458.
- Nairne, J. S. (2010). Adaptive memory: Evolutionary constraints on remembering. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 53, pp. 1-32): Academic Press.
- Nairne, J. S., & Pandeirada, J. N. S. (2008). Adaptive memory: Remembering with a stone-age brain. *Current Directions in Psychological Science*, *17*, 239-243.
- Nairne, J. S., & Pandeirada, J. N. S. (2010). Adaptive memory: Ancestral priorities and the mnemonic value of survival processing. *Cognitive Psychology*, *61*, 1-22.
- Nairne, J. S., Pandeirada, J. N. S., Gregory, K. J., & Van Arsdall, J. E. (2009). Adaptive memory: Fitness-relevance and the hunter-gatherer mind. *Psychological Science*, *20*, 740-746.
- Nairne, J. S., Pandeirada, J. N. S., & Thompson, S. R. (2008). Adaptive memory: The comparative value of survival processing. *Psychological Science*, *19*, 176-180.
- Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 263-273.

- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*, 1-25.
- Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27-53.
- Rogers, T., Kuiper, N., & Kirker, W. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677-688.
- Schulman, A. I. (1974). Memory for words recently classified. *Memory & Cognition*, *2*, 47-52.
- Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review*, *94*, 439-454.
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*, 371-394.
- Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the norms. *Journal of Memory and Language*, *50*, 289-335.
- Weinstein, Y., Bugg, J. M., & Roediger, H. L. (2008). Can the survival recall advantage be explained by the basic memory processes? *Memory & Cognition*, *36*, 913-919.

### Acknowledgements

This research was supported, in part, by a grant from the National Science Foundation (BCS-0843165).

Josefa N. S. Pandeirada was, in part, supported by the Portuguese Foundation for Science and Technology Fellowship SFRH/BPD/21534/2005.

Table 1: Rating and response time (ms) averages for survival and pleasantness (Experiment 1a) and for survival and moving (experiment 1b).

|               | Experiment 1a |              | Experiment 1b |        |
|---------------|---------------|--------------|---------------|--------|
|               | Survival      | Pleasantness | Survival      | Moving |
| Rating        | 2.91          | 3.11         | 2.79          | 2.70   |
| Response time | 3155          | 2764         | 2669          | 2591   |

Table 2: Rating and response time (ms) averages for each list in each condition from Experiment 2.

|               | Survival   |         | Robbery    |          |
|---------------|------------|---------|------------|----------|
| List          | Irrelevant | Robbery | Irrelevant | Survival |
| Rating        | 2.11       | 2.21    | 2.33       | 2.79     |
| Response time | 2429       | 2415    | 2477       | 2384     |

Table 3: Rating and response time (ms) averages for each condition in Experiment

3.

|               | Survival | Robbery |
|---------------|----------|---------|
| Rating        | 4.03     | 3.92    |
| Response time | 2340     | 2389    |

Table 4: Rating and response time (ms) averages for each list in each condition from Experiment 4.

|               | Survival   |          | Robbery    |         |
|---------------|------------|----------|------------|---------|
| List          | Irrelevant | Survival | Irrelevant | Robbery |
| Rating        | 1.98       | 4.26     | 1.90       | 4.36    |
| Response time | 2480       | 2238     | 2359       | 2224    |

*Figure 1:* The left side figure displays proportion correct recall performance for survival and pleasantness (Experiment 1a), and the figure on the right side shows proportion correct recall performance for survival and moving (Experiment 1b). Error bars represent within-subject 95% confidence intervals (as per Masson & Loftus, 2003).

*Figure 2.* Proportion correct recall performance for the irrelevant list, common to both conditions, and irrelevant lists that were scenario-dependent (survival list presented to the robbery group, and robbery list presented to the survival group) in Experiment 2. Error bars represent within-subject 95% confidence intervals (as per Masson & Loftus, 2003).

*Figure 3.* Proportion correct recall performance for the congruent lists in the survival and robbery conditions in Experiment 3. Error bars represent within-subject 95% confidence intervals (as per Masson & Loftus, 2003).

*Figure 4.* Proportion correct recall performance for the irrelevant list, common to both conditions, and the congruent lists (survival list presented to the survival group, and robbery list presented to the robbery group) in Experiment 4. Error bars represent within-subject 95% confidence intervals (as per Masson & Loftus, 2003).

Figure 1

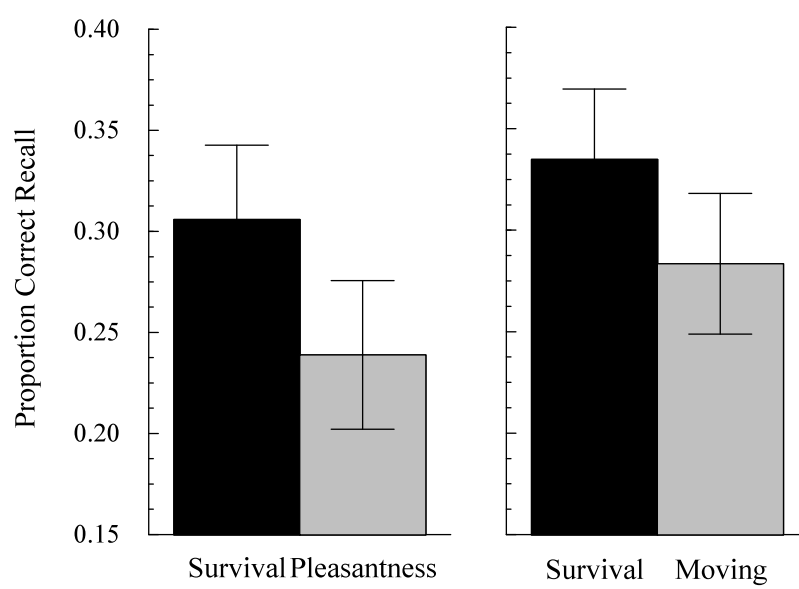


Figure 2

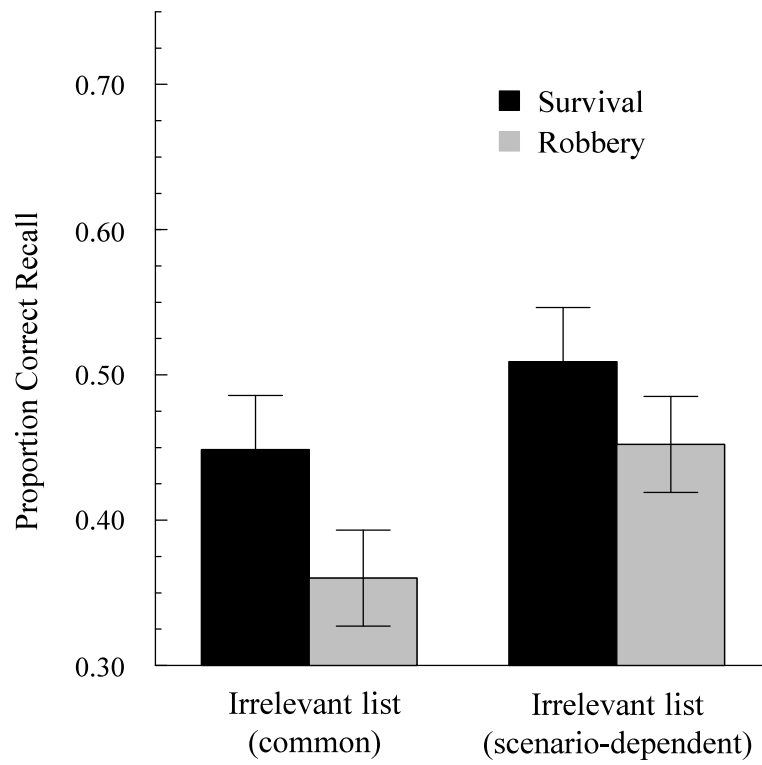


Figure 3

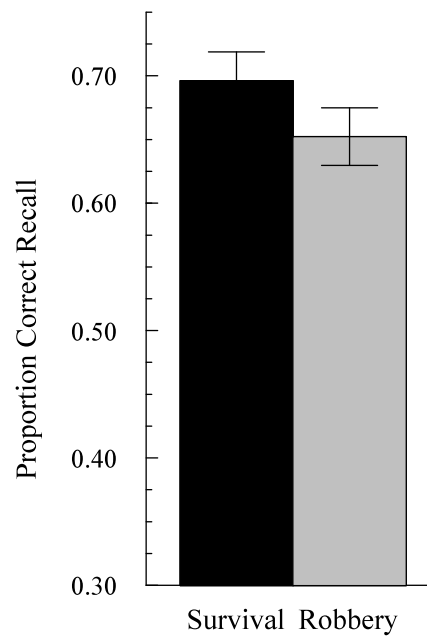


Figure 4

