Increased Meaning Elaboration
Does Not Account for the Picture Superiority Effect

A Thesis Presented to

The Faculty of Graduate Studies of Lakehead University

by

David D. C. Jeans ©

In partial fulfillment of the requirements for the degree of Master of Arts

August 2001

Thesis Supervisor: Dr. G. Hayman

0-612-60849-2

Abstract

A striking characteristic of episodic memory is that memory is better for pictures than words--the *picture superiority effect* (Paivio, 1971). While evidence in support of past explanations (e.g., dual-coding and sensory semantic models) has been inconsistent, a growing body of behavioural (e.g., Potter & Faulconer, 1975; Smith & Magee, 1980) and neurological (Grady et al., 1998) evidence points to superior processing of meaning, which is generally associated with pictorial presentations, as a major source of the pictorial superiority effect. The results of the present study--which manipulated meaning processing at study (congruent/incongruent meaning questions vs. no questions) and study/test form (picture or word at study crossed with picture or word at test)--revealed that when potential ceiling problems of previous studies (e.g., Durso & Johnson, 1980; Emmerich & Ackerman, 1979) are controlled, meaning elaboration 1) reduces, but does not eliminate, picture superiority in Yes/No Recognition responses, 2) does not affect the advantage of pictures over words in Remember (e.g., Tulving, 1985) responses but may affect Know responses, and 3) does not affect the advantage of pictures over words in Source Memory (e.g., Durso & Johnson, 1980) responses. Moreover, because the benefits of reinstating the study form at test were as large for words as for pictures--for recognition, remember, and source responses--the results imply that the processing of pictures cannot be treated as including processing in common with words with the addition of picture specific processing. Rather, the processing of pictures and words must result in equally unique sources of information that differ in terms of their overall memorability.

Acknowledgements

I would like to thank my thesis advisor, Dr. Gordan Hayman, for devoting considerable time and energy to this project. I would also like to thank the other members of my committee, Dr. Michael Wesner and Dr. Ron Davis, as well as my external reader, Dr. Joan Gay Snodgrass. Thanks to my family and friends for their support. Finally, I would like to especially thank Alissa Setliff for providing constant encouragement, support, and research expertise throughout every phase of this project.

Table of Contents

List of Tables

The picture superiority effect (Paivio, 1971)--the finding that pictures are remembered better than words--is one of the most robust phenomena in psychology and has puzzled researchers for over one hundred years (for review see Kinjo & Snodgrass, 2000; Kobayashi, 1986). It has primarily been observed on tests of episodic or explicit memory; that is, tests that require intentional recollection (Tulving, 1972; 1983; 1993). For example, numerous controlled laboratory experiments have found that when subjects try to recall or identify previously studied pictures and words, memory performance is consistently better for pictures (Weldon & Coyote, 1996). While the reason for the effect is still under active debate (e.g., Kinjo & Snodgrass, 2000; Mintzer & Snodgrass, 1999; Vaidya & Gabrieli, 2000; Weldon & Coyote, 1996), it has been suggested that pictures elicit greater elaboration of meaning than words, thus providing a "richer" or "stronger" memory trace at test (e.g., Durso & Johnson, 1979; 1980; Potter & Faulconer, 1975; Smith & Magee, 1980). Before describing an experiment in which this assumption is tested, the following paragraphs will present a brief overview of episodic memory, and discuss how factors that manipulate elaboration of meaning affect episodic memory.

Take a moment to think about what type of information is required to remember a specific event in your life; for example, the details surrounding your high school graduation. On the one hand, you would need to recall information specific to the event itself, such as what time of year it was. To do this you might want to, for example, think of how cold it was at the time, or whether or not there was snow on the ground. On the other hand, in order to recall and make sense of this specific information, you also need to have general knowledge about the world, such as the relationship between weather/temperature and time of year. In other words, to remember autobiographical events in one's life, one must retrieve and relate both specific information and general knowledge.

According to the framework developed by Tulving (1972; 1983; 1993), memory for general knowledge and specific events represent separate, but interdependent, memory systems: semantic and episodic memory, respectively. These memory systems serve to acquire, retain, and retrieve external information (Tulving 1984; 1985a). Semantic memory is composed of basic facts, ideas, rules and concepts (e.g., one's knowledge of the relationship between temperature and time of year) that are the basis of an individual's general knowledge of the world. In contrast, episodic memory refers to autobiographical events that an individual has experienced in the past; it contains knowledge of specific earlier experiences and the circumstances that surrounded them (e.g., when and where they occurred). Thus, unlike semantic memory, episodic memory codes for specific spatial locations and temporal organizations (Tulving, 1993).

Considering the complexity of spatial-temporal coding, and the similarity between the details of various life experiences, one would expect a lot of associative interference in episodic memory. For example, a person might confuse the details of his high school graduation with, say, those of his wedding. Tulving (1993) and others (e.g., Hayman, MacDonald, & Tulving, 1993; Metcalfe, Cottrell, & Mencl, 1992) hypothesize that, to deal with this complexity, episodic memory provides for the rapid cognitive binding of novel information. In other words, to avoid confusion between events composed of similar parts, episodic memory encodes components of each event *interactively*. That is, episodic memory represents information from an event in terms of the interrelations among the parts or components of semantic memory comprising that event. In this way, representations of similar events can be stored and accessed as distinct units in episodic memory.

Evidence of interactive encoding in episodic memory can be obtained by studying how the encoding of different types of information affects episodic memory performance. For example, according to Hayman, Servais and MacDonald (1999), one consequence of interactive encoding is that any increase in memory for the whole episode should also increase memory for any of its parts. That is, if episodic memory requires retrieval from an interactively encoded representation of a prior event, one would expect a positive relation between different types of memory for the same event, even when very different types of information are tested. For example, in the case of word memory, encoding both meaning (e.g., word definition) and sensory characteristics (e.g., case, rhyme or phonemic properties) should result in greater memorability for both types of information (on tests that emphasize memory for meaning and those that emphasize memory for sensory characteristics).

Superficially, this prediction is inconsistent with those of well-established theories of human memory. For example, both the *levels of processing* (Craik & Lockhart, 1972) and *transfer appropriate processing* (Morris, Bransford & Franks, 1977) approaches treat sensory and meaning information as if they are mutually exclusive[1]. That is, despite the fact that individuals are likely to encode *both* meaning and sensory characteristics in everyday life, both approaches focus on how memory performance is affected by the exclusive encoding of meaning *or* sensory characteristics; neither theory addresses how memory performance might be affected by interactive encoding. In fact, both imply that meaning and sensory information function independently to influence memory (Hayman

& Rickards, 1995). Furthermore, advocates of these theories often assume that processing information about meaning characteristics is irrelevant when the memory task requires the retention of "shallow" or sensory information (Hayman, et al., 1999). For example, Eysenck and Keane, in describing a rhyming test (which requires the identification of words rhyming with list words), stated: "What is required for this kind of test is shallow...information" (1990, p. 154).

Reflecting this "either/or" mentality is the primary method used in experiments investigating levels of processing and transfer appropriate approaches: the *single-question design* (Hayman et al., 1999). In this approach, to-be-remembered items are studied exclusively in one of two processing conditions. That is, a single orienting question which emphasizes either meaning or sensory information is presented, thus leading subjects to engage exclusively in meaning *or* sensory-based processing. For example, in an experiment that supported levels of processing, Craik and Tulving (1975) had subjects either visualize whether a word would fit in a sentence (meaning processing) or attend to typescript (sensory-based processing). They found that meaning ("deep") encoding resulted in better memory than sensory-based ("shallow") encoding. Similarly, in an experiment testing the transfer appropriate processing approach, Stein (1978) had subjects read questions emphasizing meaning (e.g., "________ has a steel blade?") or sensory (e.g., "________ has a capital "I"?") information prior to the visual presentation of a target word (e.g., "knIfe"). A crossover interaction was found between type of processing at study and test: case processing at study resulted in better case recognition than meaning processing at study, and vice versa.

Recent studies continue to use the single-question design in studying patterns of memory variability resulting from the encoding of meaning and sensory-based information. For example, Marks (1991), who examined the effect of different visual orienting tasks on the retention of pictures' names and pictorial details, had subjects encode pictures in terms of questions about category (meaning processing) or distinct physical characteristics (sensory-based processing). In support of transfer appropriate processing, recognition of picture names and picture details was superior after meaning and sensory-based processing, respectively. Again, contradicting the notions based on interactive encoding, Marks concluded that "conceptually driven [meaning] processing does not facilitate transfer on more data-driven [sensory-based] tests of recognition that emphasize...visual details" (p. 575).

According to Hayman et al. (1999) however, the design used in the above experiments is limited because it confounds necessary information with sufficient information. For the purpose of isolating information that is critical for a task, the single-question design requires that subjects either encode meaning *or* sensory characteristics. However, while information may be necessary for a task, it may not be sufficient. For example, processing the sensory characteristics of an item may be necessary for subsequent recall/recognition of these physical characteristics, but processing of meaning may be important in the likelihood of storage and retrieval of a distinctive memory of the event. In this way, attending to one type of information could aid in the retention of other, seemingly unrelated, types. However, only by crossing the processing of meaning and sensory features at study can one assess whether or not semantic processing at study modulates memory for surface information (i.e., subjects must encode to-be-remembered information in a sensory *and* a meaning context simultaneously).

This was done in a series of experiments by Hayman and colleagues (Cribbie, 1995; Hayman and Cribbie, 1998; Hayman and Rickards, 1995; Hayman et al., 1999) using a *dual-question design*. Target items were studied following the presentation of two concurrent orienting questions: one concerning meaning and one concerning sensory information. Recall that Stein (1978), using a single-question design, had subjects read questions emphasizing meaning (e.g., "________ has a steel blade?") *or* sensory (e.g., "________ has a capital "I"?") information prior to the visual presentation of a target word (e.g., "knIfe"). In contrast, Hayman et al. (1999) presented target words such as "knIfe" with both a sensory question that was congruent (e.g., "has a capital "I"?") or incongruent (e.g., "has a capital "F"?") *and* a meaning question that was congruent (e.g., "has a steel blade?") or incongruent (e.g., "is a venomous animal?"). In contrast to Stein's finding that case processing at study resulted in better memory for sensory features (i.e., case recognition) than meaning processing at study, and vice versa (supporting transfer appropriate processing), Hayman et al. (1999) found that case recognition was higher following congruent case/congruent meaning than congruent case/incongruent meaning study processing.

The results of Hayman et al.'s (1999) experiment (as well as those of Hayman and Cribbie, 1998, who replicated these findings using pictures) therefore, support the idea that a fundamental property of episodic memory is the interactive encoding of situational information. Since a memory trace is an interdependent juxtaposition of connections between various states of semantic knowledge, increases in memory for an episode will also increase memory for each of its parts--even information that is seemingly unrelated to these "parts" (in this case, meaning- versus sensory-based information) (Hayman & Rickards, 1995; Hayman et al., 1999). Thus, instead of sensory features being "separate" from meaning (as implied by the levels of processing approach) or relevant only for certain tasks (as is implied by transfer appropriate processing), interactive encoding predicts that encoding both sensory features and meaning information can result in greater memorability for both types of information.

Generation and Reality Monitoring Effects

In another series of studies, Hayman and Dew (1997) demonstrated that an interactive encoding framework might help account for variability in other examples of episodic memory seen in *generation effects* (cf. Slamecka & Graf, 1978) and *reality monitoring* (cf. Johnson, & Raye, 1981). Both of these phenomena refer to the widely-reported memory advantage of self-generated (obtained through internal processes such as reasoning, imagination, and thought) versus perceived information (obtained through perceptual processes, such as reading) (Hayman & Dew, 1997).

In the prototypical experiment demonstrating the generation effect (e.g., Slamecka & Graf, 1978), subjects were given a rule, a cue word, and the initial letter of a target word response. They were then required to self-generate the target response (e.g., with the antonym rule, subjects were given the word-letter pair, "future-P___" or cues for the target word PAST). Performance is compared with instances where subjects simply read the target word following a rule (e.g., "antonym: future-PAST"). When memory for the target word is tested, generated words are recalled or recognized better than perceived words (i.e., those that are read).

Studies have also shown that self-generation leads to better memory for the *origin* of an event (i.e., whether it was imagined or perceived), a phenomenon dubbed, the "reality monitoring effect." For example, in an experiment by Johnson, Raye, Foley, and Foley (1981), subjects were either given a first-letter cue in the context of a category name (e.g., animal - D__) and asked to generate the target (DOG) or they heard the category name spoken by the experimenter. It was found that subjects who self-generated identified the origin of the target word DOG better than when they heard "animal - DOG" spoken.

A number of explanations have been proposed to account for the memory advantage of self-generated information. Johnson et al. (1981) attribute the reality monitoring effect to "special memories" about cognitive operations. That is, they hypothesize that the effort required for self-generation results in a greater amount of information about cognitive operations, and such operations can serve as a discriminative cue about the origin of memories. Slamecka and Katsaiti (1987) argue, however, that although response generation may be a more difficult task, there is no reason to expect that such cognitive operations will result in better memory. In spite of their "superficial distinctiveness," generate and read processing are functionally the same, and the so-called "generation effect" is simply due to differential rehearsal as a function of task demands. Specifically, because studies investigating generation effects typically use mixed-list designs (i.e., generate and read items are intermixed into the same study list), subjects are drawn to engage in selective displaced rehearsal of the "generate" items (because of their "cryptic" and "fragmentary" appearance). This study-time imbalance results in an apparent generation effect. In support of their theory, Slamecka and Katsaiti found that the effect disappears when a between-list design (i.e., subjects are placed either in a "generate" or "read" condition) is used.

Begg, Snider, Foley, and Goddard (1989), however, argue that selective rehearsal is not a "general explanation" of generation effects. First, it is inconsistent with many laboratory findings. For example, Watkins and Sechler (1988) found a much larger generation effect under incidental memory conditions than under intentional conditions, even though one would expect the effect of selective rehearsal to be greater in the latter condition. Furthermore, between-subjects generation effects have been reported in tests of recognition (e.g., Slamecka & Graf, 1978). Secondly, Slamecka and Katsaiti fail to fully explain how strategic processes function to produce a generation effect (e.g., how does a "cryptic appearance" encourage rehearsal?).

Elaboration of Meaning

In contrast to the above explanations, it can be argued that (self) generation effects are a result of *elaboration of meaning*. The basis of this position is that the meaning of a concept is specified in the cognitive space by the intersection of a set of relevant dimensions (Klein & Saltz, 1976). Memory performance will depend on the extent to which relevant dimensions are activated during learning. In terms of the present discussion, generating an item may provide a structure which increases the likelihood that the target word is related to other information, thus providing a richer mnemonic basis for recognition/recall (the generation effect) and--via cognitive binding inherent in episodic memory--better linked memories for subsequent identification of study origin (the reality monitoring effect).

Hayman and Dew (1997), working within an interactive encoding framework, provide strong evidence of such a result. In their experiments, subjects self-generated a word response to a rhyme-cue word and a 2-letter stem (e.g., fast-PA__) or perceived a response in a cue-target pair (e.g., fast-PAST). To test the effect of elaboration of meaning, the experimental group also received a meaning-based orienting question that was congruent (e.g., opposite of future) or incongruent (e.g., opposite of bright) with the target response (e.g., PAST). Hayman and Dew found that congruent questions increased recognition and reality monitoring accuracy (i.e., identifying the origin of the study words as "read" or "solved") for perceived responses to that of self-generated responses. In contrast, incongruent questions decreased performance for self-generated responses to the level of perceived responses. The results showed that 1) encouraging meaning elaboration (via congruent meaning-based questions) facilitated reality monitoring and recognition as much as the self-generation of information, and 2) disrupting meaning elaboration (via the presentation of incongruent meaning-based questions) negated any potential advantage of self-generation.

Hayman and Dew (1997) concluded that their results supported the interactive encoding approach. That is, as predicted by interactive encoding, a manipulation that increased the overall availability of an episode (i.e., elaboration of *meaning)* improved memory for a separate component of the episode (i.e., reality monitoring, a task which depends primarily on the availability of *sensory* features) (see Johnson & Raye, 1981; Marks, 1991). In addition, Hayman and Dew's results suggest that greater meaning elaboration during self-generation tasks underlies the facilitation observed in both the reality monitoring and the generation effects. That is, rather than being due to "special memories" (Johnson et al., 1981) or selective displaced rehearsal (Slamecka & Katsaiti, 1987), facilitation from imagine (generate) processing at study is thought to reflect greater attention to meaning, which provides a richer cognitive framework for binding in episodic memory.

The Picture Superiority Effect

The success of Hayman and Dew (1997) in accounting for the memory advantage of (self) generation raises the possibility that differences in meaning elaboration could also account for other examples of unexplained episodic memory variability. For example, the *picture superiority effect* (Paivio, 1971)--the finding that pictures are remembered better than words on tests of free recall and recognition--has been investigated in numerous studies (e.g., Bajo, 1988; D'Agostino, O'Neill, & Paivio, 1977; Durso & Johnson, 1979, 1980; Intraub & Nicklos, 1985; Job, Rumiati, & Lotto, 1992; Kinjo & Snodgrass, 2000; Kobayashi, 1986; Madigan, 1983; Mintzer & Snodgrass, 1999; Nelson, 1979; Nelson, Reed, & McEvoy, 1977a; 1977b; Nelson, Reed & Walling, 1976; Paivio, 1971, 1986, 1991; Paivio & Csapo, 1973; Ritchey, 1980; Smith & Magee, 1980; Snodgrass & McLure, 1975; Vaidya & Gabrieli, 2000; Weldon & Coyote, 1996). However, reasons for the effect are still unclear.

One of the earliest explanations for the *picture* superiority effect was Paivio's (1971; 1986; 1991) *dual-coding* hypothesis. Paivio proposed that imaginal and verbal representations are stored in functionally independent, but interconnected, systems in long term memory. Because encoding redundancy presumably improves memorability, Paivio suggested that events represented with both codes are more likely to be remembered than events represented with a single code. Regarding the *picture* superiority effect, Paivio argued that individuals are more likely to spontaneously name pictures than to imagine word referents (i.e., pictures are more likely to be dually encoded). Thus, they are remembered better than words. Support for the theory is provided by Paivio and Csapo (1973), who investigated the relative contributions of imaginal and verbal codes in picture/word memory. Using orienting tasks which required subjects to encode stimuli either verbally (by writing or pronouncing the words and picture labels) or imaginally (by drawing or imagining the pictures and labels), they found that recall for pictures was better than words under all conditions except when subjects imaged words. That is, when subjects were led to dually encode words, their memory for words was equal to that of pictures.

An alternative explanation of the *picture* superiority effect was provided by Nelson (1979). His *sensory-semantic* model implicates the role of stimulus surface features in superior picture memory; that is, pictures provide more distinctive visual representations than do words, thus providing a more differentiating mnemonic than the sensory representation provided by its label. This explanation is supported by evidence that high visual similarity among pictures eliminates or reverses the picture superiority effect (Nelson et al., 1976).

While there is some support for both the dual-coding and sensory-semantic approaches, both theories have been challenged by contradictory findings. For example, contrary to Paivio's (1971) hypothesis that verbally labeling pictures results in a stronger memory trace (because of dual-coding), Intraub (1979), in an investigation of implicit naming in pictorial encoding, found no correlation between naming latency and memory. Job et al. (1992), on the other hand, reported evidence contradicting Nelson's (1979) sensory-semantic model. They found that categorization of pictures *and* words was slower when the to-be-categorized stimuli were from two visually similar categories (i.e., fruits-vegetables pictures/words) than from visually dissimilar categories (i.e., fruits-weapons pictures/words). Based on this result, Job et al. argued that the *picture* superiority effect in categorization is not due to the distinctive sensory features of pictures over words. Instead, they suggested that their results could be explained most parsimoniously with reference to the effect of *semantic* (rather than visual) similarity.

This explanation is consistent with theoretical orientations which stress the importance of meaning processing (rather than sensory-based processing, e.g., Nelson, 1979; Paivio, 1971) in producing the picture superiority effect. This viewpoint is based on findings that pictures access meaning information more readily than do words (Weldon & Coyote, 1996). For example, Potter and Faulconer (1975) found that naming a pictured object (which requires phonemic or sensory information) took much longer than naming a word, but deciding whether an object was in a given category (which requires meaning information) took less time for pictures than words--a finding which was replicated by Pellegrino, Rosinski, Chiesi, & Siegel (1977). Similarly, Smith and Magee (1980) observed that word categorization was disrupted by the presence of a picture from an incongruent category, but picture categorization was not disrupted by a category-incongruent word. Based on this result, they suggested that meaning is accessed faster from pictures than words, since the incompatible information from the distractor would interfere only if that information were available prior to the production of the target response. Providing more evidence that pictures are more likely to access meaning, Nelson et al. (1977a), using a serial learning task, found that high semantic similarity (e.g., all stimuli were animals) disrupted the learning of pictures (i.e., eliminated the typical picture superiority effect) but not words.

Further evidence that pictures access meaning more readily than words is provided by studies in which orienting tasks have been used. For example, numerous studies have found that semantic orienting tasks yield better recognition and recall for words (e.g., Craik & Tulving, 1975; Geis & Hall, 1976; Hyde & Jenkins, 1973; Moscovitch & Craik, 1976) but not for pictures (e.g., D'Agostino et al., 1977; Durso & Johnson, 1980; Emmerich & Ackerman, 1979; Intraub & Nicklos, 1985). Intraub and Nicklos argue that meaning elaboration is not an automatic process for words, and tasks that encourage attention to richer meaning will improve word memory. Semantic processing of pictures, on the other hand, may be redundant with automatically initiated processes, resulting in little memory change.

Evidence for the actual "source" of these "automatically initiated" processes is provided by data from neuroimaging studies, which have revealed differences in the functional neuroanatomy for picture and word memory. Results from these studies suggest that pictures and words employ similar brain resources during semantic, but not nonsemantic, encoding. For example, Grady, McIntosh, Rajah & Craik (1998) found that picture encoding resulted in more extrastriate cortex and medial temporal cortex activation than word encoding under "neutral" conditions (see also Menard, Kosslyn, Thomson, Alpert, & Rauch, 1996; Nyberg, 1999). However, when pictures and words were presented with semantic orienting questions (which presumably encouraged similar levels of picture/word meaning elaboration), no picture/word differences in medial temporal cortex activation were found (Grady et al., 1998).

Grady et al. (1998) noted that the extrastriate cortex is generally activated during visual perception of *both* verbal and nonverbal material (cf. Haxby, Grady, Horwitz, Ungerleider, Mishkin, Carson, Herscovitch, Schapiro, & Rapoport, 1991; Peterson, Fox, Snyder & Raichle, 1990; Zeki, Watson, Lueck, Friston, Kennard, & Frackowiak, 1991) and was probably more active during picture encoding in their experiment because the pictures used were more visually complex than the words. The medial temporal cortex, on the other hand, has long been known from lesion experiments to be important for episodic memory (Aggleton, Hunt, & Rawlins, 1986; Mishkin, 1978; Scoville & Milner, 1957; Sutherland & McDonald, 1990), and may be particularly important for encoding new information (Squire, 1992). The presence of differential activation in this area during picture/word processing suggests that it may also be the "source" of picture superiority. Furthermore, the fact that this differential activation disappeared during semantic encoding suggests that meaning elaboration may be the mechanism underlying the effect, and conditions that equate picture/word meaning elaboration may also equate picture/word memory. Indeed, the behavioural results of Grady et al.'s study support this hypothesis: performance on recognition tests showed similar levels of incidental memory for pictures and words (i.e., a reduction of the picture superiority effect) following semantic processing, but not nonsemantic processing (which showed the usual picture superiority effect).

<u>Methodological Problems of Past Studies</u>

While Grady et al. (1998) found that picture/word memory was similar following meaning elaboration, they could not make strong conclusions regarding this trend because they only tested enough subjects to achieve statistically significant neurological (rather than behavioural) results. Unfortunately, other studies examining the effect of meaning elaboration (via orienting questions) on picture/word memory have also been inconclusive due to methodological limitations. For example, Emmerich and Ackerman (1979), tested picture/word recognition and response latencies following "acoustic" (e.g., "Does it rhyme with kale?" for NAIL), "schematic" (e.g., "Is it round?" for SUN), and "conceptual" (e.g., "Does it unlock things?" for KEY) orienting questions. While their results showed a trend towards equivalent picture/word memory performance (i.e., an elimination of the picture superiority effect) following conceptual questions (rather than acoustic and schematic questions), they only reported analyses for their response latency data because of near ceiling levels of recognition performance (i.e., nearly all pictures and words were recognized at test). The fact that ceiling effects were present for pictures and words in the "conceptual" questions condition (91% and 93% recognition for pictures and words, respectively), but only for pictures in the control condition (90% recognition, as compared to 82% for words), suggests that additional benefit to pictures due to meaning elaboration in the "conceptual" questions condition was not detected. Therefore, Emmerich and Ackerman could not conclude that meaning elaboration eliminated the picture superiority effect in their experiment.

Other studies have found that meaning elaboration appeared to eliminate the picture superiority effect, but methodological problems have left open the possibility that results were due to confounding variables. For example, D'Agostino et al. (1977), who examined picture/word recall following the presentation structural, phonemic, and semantic orienting tasks, obtained significant results which were somewhat contradictory to the present hypothesis: Semantic *and* phonemic processing produced equivalent picture/word recall performance. However, as Durso and Johnson (1980) argue, D'Agostino et al.'s experiment was flawed because the test stimuli and orienting questions were presented twice. Repetition of the concepts would tend to benefit the weaker trace relative to the stronger one and could have resulted in equivalent recall, not because picture and word traces were similar, rather, because words accrued the benefits of repetition more rapidly (cf. Durso & Johnson, 1979).

In an attempt to correct the confound in D'Agostino et al.'s (1977) experiment, Durso and Johnson (1980) conducted a study examining picture/word recognition following single-presentations of "referential", "imaginal" and "verbal" orienting tasks (which emphasized semantic, visual, and phonemic information, respectively). They reported an elimination of the picture superiority effect in recognition following presentations of "function" questions (i.e., "What is the object used for?"), and a reversal of the picture superiority effect (i.e., a word superiority effect) following presentations of "explicit imagery" questions (i.e., "Create an image"). However, in addition to the presence of ceiling effects in the above conditions (mean recognition of 93% for pictures and words), Durso and Johnson's (1980) experiment contained another critical confound: the absence of a control for study/test form interactions (i.e., they tested recognition of studied pictures and words at study using only aurally-presented words). Such a procedure is problematic for two reasons. First, absence of a test form manipulation limits the detection of picture/word memory differences to study only (i.e., picture/word differences in both encoding and retrieval processes cannot be assessed). Several researchers have argued that retrieval processes are often neglected in studies of the picture superiority effect, despite recent data suggesting that they play an important role (e.g., Mintzer & Snodgrass, 1999; Nicolas, 1995; Sternberg, Radeborg, & Hedman, 1995; Weldon & Coyote, 1996; Wippich, Melzer, & Mecklenbrauker, 1998).

Second, and more importantly, one of psychology's most consistent findings is that, as dictated by the encoding specificity (Tulving & Thomson, 1973) and transfer appropriate processing (Morris et al., 1977; Weldon, Roediger & Challis, 1989) theories, memory is better when there is a match between cues at study and test. In studies like Durso and Johnson's (1980), pictures at study are being tested with a different form of test stimuli, while words at study are tested with the same form of test stimuli. Therefore, this procedure theoretically presents a bias favouring word (rather than picture) retrieval.

Direct evidence supporting this possibility is provided by studies which have measured the effect of changing stimulus form (picture or word) between study and test. For example, Sternberg, et al. (1995), in a study investigating the effect of form-changing on recognition memory reaction time (RT), found that a larger decrement in performance was obtained when pictures were studied as words than when words were studied as pictures (i.e., there were larger "form-change costs" for pictures than words). Mintzer and Snodgrass (1999) replicated this finding in an experiment testing recognition memory. These studies suggest that Durso and Johnson's design could have eliminated picture superiority with or without meaning elaboration. As Mintzer and Snodgrass suggest, "the advantage of the same form word-word item over the different form picture-word item could be counteracted by the advantage of a picture at study over a word at study, producing approximately equal performance for the word-word and picture-word items" (p. 122). This is precisely the result that Durso and Johnson obtained.

The above analysis may also apply to studies that show a reduction, elimination, or reversal of the picture superiority effect following meaning elaboration in recall (D'Agostino et al., 1977; Durso & Johnson, 1980). For example, Durso and Johnson (1980) found that "referential" and "imaginal" orienting questions also eliminated the picture superiority effect in recall. However, recall tests can also present a study/test form interaction that might counteract the picture superiority effect because it involves the self-generation of retrieval cues. Specifically, if one is inclined to self-generate verbal rather than pictorial cues, there would be a match between encoding and retrieval cues for words rather than pictures that would counteract picture superiority in a manner similar to the word-word condition of a recognition test.

Another weakness of past studies examining the effect of meaning elaboration on picture/word memory is their failure to measure memory following the presentation of semantically *incongruent* orienting tasks. Recall that Hayman and Dew (1997), in their examination of self-generation effects, found that disrupting meaning elaboration (via the presentation of incongruent meaning-based orienting questions) negated any potential advantage of self-generation. Analogously, if superior meaning elaboration is the reason for the picture superiority effect, one would expect that manipulations which disrupt such meaning elaboration would also reduce this picture/word memory difference. While Hayman and Cribbie (1998) reported that incongruent meaning questions did result in poorer memory for pictures than congruent meaning questions, to our knowledge, no such manipulation has ever been conducted in the context of comparing picture and word memory. Comparisons of picture/word memory under congruent meaning, incongruent meaning and control conditions would presumably shed much light on the current debate.

Future studies investigating the effect of meaning elaboration on the picture superiority effect would also be better served by using materials that are thoroughly evaluated and normed. Past studies have used pictures that were hand drawn by the authors using a felt tip pen and not extensively judged for name or image agreement (e.g., Durso & Johnson, 1980; Emmerich & Ackerman, 1979). It is therefore possible that they were less recognizable or led to smaller picture superiority effects. A smaller picture superiority effect would, in turn, be more easily masked or hidden by other effects, such as the study/test form interaction discussed earlier.

In summary therefore, the debate is still open as to whether elaboration of meaning can eliminate the picture superiority effect and, more importantly, whether differences in meaning elaboration, on the whole, can account for picture/word memory differences. Studies investigating the former question using implicit tests of memory have also been conducted. Evidence supporting (e.g., Nicholas, 1995; Weldon & Roediger, 1987; Weldon, Roediger, & Challis, 1989; Wippich et al., 1998) and negating (e.g., Weldon & Coyote, 1996) the hypothesis have been found. While these studies are interesting, results using implicit tests of memory lie outside our domain of interest for two reasons. First, numerous variables have been found to influence explicit but not implicit memory performance, such as levels of processing (Graf & Mandler, 1984; Jacoby & Dallas, 1981), generation versus read study conditions (Gardiner, 1988b), divided versus undivided attention (Parkin, Reid, & Russo, 1990), and intentional versus incidental learning (Greene, 1986). Because of the apparent differences between explicit and implicit memory, direct comparison of results based on them is unwise. Secondly, we are interested in what our results suggest about the nature of episodic memory. As

Tulving (1983) asserted: "In theories of episodic memory, recollective experience should be the ultimate object of interest, the central aspect of remembering that is to be explained and understood" (p. 184). Implicit memory tests, in contrast, are by definition, those in which the "conscious recollection of prior events and experiences is not required" (p. 617) (Gardiner, 1988a), and thus are incompatible with the focus of our study.

<u>Remember/Know Tests of Memory</u>

Consistent with our focus on recollective experience, the present study will also make use of a technique suggested by Tulving (1985b) for measuring the nature of subjects' conscious awareness during tests of memory. Tulving (1985b) described experiments in which subjects were required to put an "R" (for "remember") next to items in the test whose prior occurrence in the study list they could consciously recollect, and a "K" (for "know") next to items they recalled or recognized on some other basis. Within Tulving's framework (1983, 1985a, 1985b), "remember" responses reflect output from episodic memory, because recollective (or autonoetic) consciousness is a defining characteristic of that system. A "know" response (or noetic consciousness), on the other hand, is characteristic of semantic memory, because knowledge retrieved from semantic memory is not normally accompanied by recollective experience (Gardiner & Java, 1991; 1993; Rajaram, 1993; 1996). This conceptualization is supported by evidence suggesting a dissociation between remember/know phenomena--similar to dissociations observed in studies that have manipulated conscious awareness by comparing performance in explicit and implicit memory tests (Gardiner, 1988a). For example, variables such as levels of processing and generate-versus-read study conditions (Gardiner, 1998a), divided-versus-undivided attention (Gardiner & Parkin, 1990), and intentional-versus-incidental learning (Macken & Hampson, 1991) have been found to be influenced by "remember," but not "know," responses. Remember/know recognition tests therefore, are a useful addition to studies investigating principled outcomes of episodic memory manipulations.

Present Study

The purpose of the present study was 1) to examine the extent that differences in meaning elaboration during study accounts for differences in picture/word memory, and 2) to investigate interactive encoding in incidental episodic memory. Memory for pictures and words was assessed in two instruction groups: a no_questions group that viewed pictures and words under neutral conditions (i.e., following the presentation of a series of "*" symbols), and a questions group that was required to answer congruent/incongruent questions about the meaning of the presented stimuli. To avoid the limitations of past studies, our experimental design included, in addition to the manipulation of meaning processing at study (congruent/incongruent questions versus no questions), a 1) factorial manipulation of study and test form (picture or word), 2) a filler task and several recency buffers, presented between the study and test conditions, to avoid ceiling effects, and 3) picture and word materials taken from Snodgrass & Vanderwart (1980), which have been thoroughly normed and standardized. Memory was measured using a remember/know test of study form. That is, for each picture/word presented at test, subjects indicated *what* they recollected (picture or word study form), and *how* they recollected it ("remember," "know," "not at all"). Responses on the remember/know test were used to create four dependent variables (recognition, remember, know, and correct source identification responses), so that we could assess the effect of our manipulations on various indices of memory.

It was expected that there would be an overall picture superiority effect in the present study that would interact with test form. That is, picture superiority was expected to be larger when pictures, rather than words, were presented at test. In terms of the effect of the manipulation of meaning, two hypotheses were presented. First, it was hypothesized that increased elaboration of meaning at study would facilitate memory for words more than for pictures; that is, the picture superiority effect would be reduced in the questions group only with congruent questions. Second, it was hypothesized that incongruent meaning questions would disrupt meaning encoding more for pictures than for words; that is, memory performance with pictures was predicted to be lower with incongruent questions in the questions condition than in the no_questions condition; it was expected that memory performance with words would be similar in the incongruent question and no_questions conditions.

In terms of the differences between the dependent variables (i.e., remember, know, source identification), certain patterns are expected. First, because source memory is, by definition, "the ability to remember the context in which a particular piece of information has been learned", (Rybash et al., 1995, p. 112), it has been argued that the processes underlying source and remember judgments are neurally and functionally equivalent (e.g., Conway & Dewhurst, 1994; Dewhurst & Conway, 1994; Dewhurst & Hitch, 1999; Rugg, Schloerscheidt, & Mark, 1998; Wilding & Rugg, 1996). This notion has been supported by neuroimaging studies showing that the electrophysiology of remember and source responses are similar (Rugg et al., 1998). Thus, in the present experiment, it was expected that source identification and remember responses would show a similar pattern of results. Second, as mentioned above, remember and know judgements are thought to represent recollection (based on the conscious retrieval of a specific study episode) and familiarity (based on information devoid of contextual context), respectively (Gardiner & Java, 1991; 1993; Rajaram, 1993; 1996; Tulving, 1983; 1985a; 1985b). In support of this view, several variables have been shown to affect remember and know responses in different ways, including Rajaram's (1993) finding of opposite picture/word memory performance in remember and observed know responses (i.e., she found picture superiority in remember responses and word superiority in observed know responses). Therefore, it is expected than remember and know results will be differentially affected the by manipulations in the present experiment, providing further evidence for a dissociation between remember and know phenomena. In particular, based on Rajaram's (1993) result, we expected picture superiority effects to be observed in remember, but not know, results.

A subsidiary purpose of the present design was to investigate the interdependent nature of episodic memory between incidentally and intentionally encoded events. For this purpose, the experiment included a manipulation of sensory features (colour) at study and test. Recall that the notion of interactive encoding suggests that when memory for any part of the episode is improved (e.g., memory for meaning), there should be positive benefits in memory for other parts of the episode, even otherwise unrelated information (e.g., memory for sensory information) (Hayman & Rickards, 1995). Hayman and Cribbie (1998) found that memory for task relevant sensory features (i.e., picture colour) was higher following congruent meaning and incongruent colour processing than incongruent meaning/congruent colour, congruent meaning/incongruent colour, and incongruent meaning/incongruent colour processing. Our interpretation of interactive encoding suggests that memory for irrelevant information (i.e., study colour when no orienting questions about sensory information is presented at study) should be similarly affected by manipulations of meaning, to the extent that some degree of incidental colour encoding occurs. Therefore, while the present study manipulated explicit attention only to meaning, it was predicted that, in a 4AFC (four-alternative forced choice) test of memory for the colour of the study stimulus, memory for sensory features (colour) would be better for pictures, and perhaps words, studied following congruent rather than incongruent questions. Previously, Kolers, Duchnicky, and Sundstroem (1985) found that memory for pictures was more sensitive to variation in sensory features between study and test than was memory for the names (words) of these pictures. Similarly, we expected incidental memory of colour to be better for pictures than for words. Finally, because interactive encoding applies to episodic and not semantic memory (Hayman et al., 1993; Metcalfe et al., 1992; Tulving, 1993), this facilitation in sensory memory from meaning elaboration is expected to depend on remember/know recognition. That is, better memory for sensory features was expected following congruent than incongruent questions only for items which are represented in and retrieved from episodic memory (i.e., those rated "remember" rather than "know").

Method

Subjects

Seventy-two introductory psychology students (30 males and 42 females) at Lakehead University participated in the experiment for course credit. They were pseudo-randomly assigned to one of four groups.

Materials

One hundred and eighty four pictures (black on white line drawings) and their verbal labels were selected from those used by Snodgrass and Vanderwart (1980). One hundred and twenty were used as targets, 24 as study buffers, and 40 as recognition test lures. These stimuli are standardized according to four variables related to processing: 1) name agreement, 2) image agreement, 3) familiarity, and 4) visual complexity (Snodgrass & Vanderwart, 1980).

Of the 16 categories in Snodgrass and Vanderwart's original set, ten (animals, birds, clothing, fish, furniture, grooming, insects, kitchen supplies, miscellaneous, and musical instruments) were used to construct lists of target pictures/words. Six sublists of 20 target pictures/words each were constructed, using items from each of the 10 categories in each sublist. For each subject, four of the six sublists were presented during study and two were presented during the test phase (as nonstudied test lures). To counterbalance the presentation of target pictures/words at study and test, the sublists were rotated among the subjects, such that each of the six sublists appeared in each study and test condition an equal number of times.

The 24 study buffers and 40 recognition lures were selected from the remaining six categories (tools, toys, vehicles, body parts, fruits, and vegetables) of Snodgrass and

Vanderwart's (1980) set. The study buffers were used to control for primacy and recency effects. Eight primacy items and 16 recency items were presented at the beginning and end of the study list. Thus, a total number of 104 items (80 target items and 24 study buffers) were presented during the study phase.

To equate the proportion of "old" and "new" items presented during the remember/know test phase (and therefore reduce the likelihood of a response bias), 40 recognition lures were included. These items were never present at study (i.e., they were nonstudied items for all subjects). Because these items were different material from the targets, responses were recorded but not analyzed. Thus, combined with the 40 items from the two nonstudied target sublists (20 from each of 2 sublists), the remember/know test consisted of 80 nonstudied items (40 items from two target sublists and 40 recognition lures) and 80 studied items (all from the four target sublists).

To test memory for sensory features (i.e., colour), the black and white pictures from Snodgrass and Vanderwart's (1980) original set were transformed into coloured line drawings on a black background. Each picture was assigned 4 different colours from a pool of 8 colours (red, blue, green, yellow, orange, brown, pink and purple). Colours were assigned with the constraint that: 1) no picture was assigned a colour in which a prior association existed (e.g., a picture of a frog would not be assigned the colour green), and 2) each of the eight colours occurred with roughly equal frequency. Therefore, chance performance in a 4AFC test for nonstudied pictures should be at or close to 25% if the target colour and the test lure colours had no prior association with the pictures. This assumption was tested using the nonstudied targets or lures.

To manipulate meaning, each picture was assigned a congruent or incongruent description (in question form) of it's meaning (only one of which was presented at study). The congruent questions made reference to a defining characteristic of the object depicted in each picture. For example, included were statements regarding the function or typical uses of the object (e.g., "used to kiss?" for lips), popular beliefs about the object (e.g., "keeps the doctor away?" for apple), object animacy (e.g., "back and forth?" for a swing), and physical descriptions ("a spotted cat?" for a leopard). Alternatively, incongruent questions assigned to each picture were not related to the meaning of the object depicted (e.g., "a sticky condiment?" for a truck). Care was taken to ensure that the respective congruent/incongruent descriptions for each picture were unrelated. For example, the congruent question about the picture of a horse was "a cowboy's transportation?" whereas the incongruent meaning statement was a "dangerous substance?" (rather than a reference to a similarly phrased, but opposite, concept). A Macintosh Power 6400 computer was used to present the stimuli, and to collect and tabulate responses.

Design

The design of the experiment was a mixed factorial (between/within). The between-subjects factors were: 1) instructions (questions or no_questions) and 2) test form (pictures or words at test). The within-subjects factors were: 1) study form (pictures or words at study) and 2) question context (congruent or incongruent). There were two measures: 1) responses on the remember/know test (remember, know, or nonstudied), and 2) responses on the 4AFC test of colour recognition.

Procedure

Subjects were tested individually. To ensure that the colours could accurately be distinguished, the eight possible colours were first presented on the computer screen along with their respective labels (i.e., red, blue, green, yellow, orange, brown, pink and purple), and subjects were asked to distinguish between them. In the questions group, a congruent/incongruent question was presented on the computer screen for four seconds; in the no_questions group a series of "*" symbols were presented. Following a one second blank screen, a coloured picture or word was presented for three seconds, accompanied by a "query box" (that remained until the subjects responded) containing the options, "1) Yes" and "2) No." The subjects were required to respond (by pressing "1" for "yes" and "2" for "no") as to whether the preceding phrase accurately described the picture or word (i.e., whether it's meaning was "congruent" or "incongruent"). In the no_questions group, the query box contained the words, "1) To Continue" (since a congruent/incongruent decision was not required). Subjects in the no_questions group responded by pressing "1." Subjects were given as much time as necessary to respond. The next trial began two seconds after subjects responded.

Immediately following the Study Phase, each subject was asked to complete a paper and pencil "famous names" quiz for 10 minutes. This was done in order to reduce the likelihood of ceiling effects on the following tests (Cribbie, 1995; Rajaram, 1993). In the quiz, subjects were required to answer a series of multiple-choice questions pertaining to celebrities from a variety of domains (e.g., well-known authors, actors/actresses, musicians, etc.).

Following the "famous names" quiz, each subject was first asked to read descriptions of "remember" versus "know" recognition (see Appendix A). Once subjects had finished reading and indicated that they understood the difference between "remember" and "know" judgments, the experimenter verbally repeated the distinction and answered any questions. Subjects were then presented with a remember/know recognition test containing a total of 160 pictures or 160 words (80 study targets, 40 nonstudied targets and 40 test lures). The subjects responded by selecting one of five options: 1) Remember Picture from study session, 2) Remember Word from study session, 3) Know Picture (or similar one) was studied, 4) Know Word (or similar one) was studied, or 5) Not studied as picture or word. There was no time limit to respond. Subjects responded by using the keyboard, after which the next trial was presented (following a one second blank screen).

Following the remember/know recognition test, the 4AFC test was presented: subjects were given a 4AFC test of 120 pictures or 120 words (80 previously-studied target items and 40 nonstudied target items, presented individually). Each picture and word was presented simultaneously in four different coloured versions, one in each of the four corners of the computer screen. For studied targets, one colour matched the study colour; for lures, subjects had to guess. When one of four designated responses (the keys "1", "2", "4" or "5") were pressed on the keypad, an arrow would appear in the centre of the screen, pointing in the direction of one of the 4 coloured-versions of the picture or word (i.e., the numbers "1", "2", "4", and "5" pointed the arrow at corresponding quadrants of the screen). Subjects were required to choose the colour in which they believed the study items were presented in Study Phase. They were instructed to guess if they were not sure of the answer.

Following the experiment, subjects were debriefed as to the nature and purpose of the study. They were also instructed on whom to contact regarding any concerns about the experiment.

Results

Remember/Know Recognition Test

Subjects' responses from categories one to five (i.e., remember picture, remember word, know picture, know word and not studied) were used to create four dependent variables: recognition, remember, know, and correct source responses for studied and nonstudied pictures and words. Recognition responses were computed by summing across categories one to four, inclusive. Remember responses were computed by summing across response categories one and two. Know responses were computed by summing across response categories three and four. Correct source responses were computed by summing across response categories one and three for pictures, and two and four for words. The effect of the experimental manipulations on these four dependent variables will be considered in turn.

Recognition Responses

Table 1 displays the mean proportion of items correctly recognized in all 40 test conditions--four between-subject (two instructions X two test form) by ten within-subject (two study form X two question context X two colour, plus nonstudied items) conditions. By visual inspection, there appeared to be no effect of "same" (rows one to four) versus "different" (rows five to eight) study/test colour. A preliminary analysis (involving the

32 conditions) looking at the benefit of maintaining the study colour of tested pictures and words confirmed this: the main effect ($F < 1$) and all interactions involving colour were not significant. The main effect of colour was [$F(1, 68) = .836$, $MS_e = .013$] and the two-way interactions involving colour were [$F(1, 68) = 2.25$, $MS_e = .097$; $F(1, 68), = .066$, $MS_e = .097$; $F(1, 68), = .096$, $MS_e = .015$; and $F(1, 68), = 1.88$, $MS_e = .014$ for instructions, test form, study form, and question context, respectively (all $ps > .050$)] for recognition responses. Interactions involving test colour which failed to achieve significance were: a three-way interaction between test form and question context [$F(1, 68) = 1.52$, $MS_e = .019$], a three-way interaction between study form and question context [$F(1, 68) = 2.18$, $MS_e = .013$], a four-way interaction between instructions, study form, and question context [$F(1, 68) = 2.62$, $MS_e = .013$], and a five-way interaction between instructions, test form, study form, and question context [$F(1, 68) = 3.11$, $MS_e = .013$] (all $ps > .100$). All other (three-, four-, and five-way) interactions involving colour were not significant (all F's < 1). Thus, the data for responses in "same" and "different" colour conditions were collapsed in the remaining recognition analyses to simplify interpretation.

Recognition hits – collapsed over colour.

Table 2 displays the mean proportion recognition, collapsed over colour, for the four between- (two instructions X two test form) and four within-subject (two study form X two question context) conditions. Inspection of the means revealed an overall advantage of study form, pictures greater than words, although the magnitude of the study form advantage appeared to interact with the instructions group manipulation (see rows five and six). This impression was confirmed in a 2 X 2 X 2 X 2 mixed-factors

MANOVA, with instructions (questions group versus no_questions group) and test form (picture versus word) as between-subject factors, and question context (congruent versus incongruent) and study form (picture versus word) as within-subject factors. There were significant main effects of instructions [$F(1, 68) = 18.46$, $MS_e = .048$, $p < .001$] and study form [$F(1, 68) = 138.68$, $MS_e = .017$, $p < .001$], and a significant interaction between these variables [$F(1, 68) = 12.18$, $MS_e = .017$, $p < .001$]. Differences between means leading to these main effects and interactions were evaluated in planned comparisons within groups using Fisher's LSD of .061 derived from the MS_e term of the within-subject factor interaction. The advantage of study picture over study word was larger in the no_questions (means of .83 and .60 for pictures and words, respectively, for a difference of .23) than in the questions group (means of .89 and .76, for a difference of .13). An alternative interpretation of the previous interaction is to consider how study form affected the between-group effects of instructions. A between-group comparison using Fisher's LSD of .103 (derived from the MS_e term of the between-subject factor of instructions) found that the addition of orienting questions significantly increased recognition of words (questions and no_questions group means of .76 and .60, respectively, for a difference of .16), but not of pictures (means of .89 and .83, for a difference of .06).

Although the three-way interaction between instructions, study form and question context was not significant [$F(1, 68) = 11.02$, $MS_e = .007$, $p = .067$], planned comparisons within groups (Fisher's LSD of .039) were nonetheless conducted because it had been hypothesized that picture superiority would be reduced in the questions group only with *congruent* orienting questions. While the advantage of study picture over word was smaller in the congruent questions condition (difference of .08) than in the incongruent questions condition--which was the same as in the congruent/no_questions and incongruent/no_questions conditions (differences of .23, .25, and .17, respectively)--responses to pictures were significantly greater than to words in all conditions. However, this apparent reduction of the picture/word advantage in the congruent questions condition may be due to ceiling effects, as will be addressed in the discussion.

Inspection of the means in Table 2 also revealed that, as predicted, the benefit of question context occurred in the questions group (row six), rather than in the no_questions group (row five) because subjects in the no_questions group were presented with "******" rather than congruent or incongruent orienting questions. This was confirmed by both a significant main effect of question context [$F(1, 68) = 55.58$, $MS_e = .013$, $p < .001$], and a significant question context X instruction group interaction [$F(1, 68) = 40.31$, $MS_e = 5.00$, $p < .001$]. Planned comparisons within groups (Fisher's LSD = .053) found that the benefit of congruent over incongruent context was significant in the questions group (difference = .180) but not the no_questions condition (difference = .015).

In addition to instructions, study form also interacted with test form [$F(1, 68) = 26.59$, $MS_e = .017$, $p < .001$] and question context [$F(1, 68) = 8.73$, $MS_e = .007$, $p < .001$]. There was also a significant three-way interaction between study form, test form, and question context [$F(1, 68) = 11.02$, $MS_e = .007$, $p < .001$]. Post hoc comparisons (Fisher's LSD = .039) suggest that the study picture over study word advantage was significant in each comparison, but was largest for test pictures, especially in the incongruent condition (difference of .32, as compared to .20, .11 and .10 in the congruent/test picture, congruent/test word and congruent/test picture conditions, respectively).

The main effect of test form failed to achieve significance [$F(1, 68) = 1.01$, $MS_e = .048$, $p > 1$], as did its interactions with instructions and question context (both Fs < 1). All other (two-, three-, and four-way) interactions were not significant (all Fs < 1).

In summary, as predicted, recognition responses showed an overall advantage of study pictures over study words that interacted with instructions and question context. There was support for the hypothesis that picture superiority would be reduced by congruent orienting questions, in that the advantage of study pictures over study words was smallest in the congruent questions condition.

Recognition false alarms – collapsed over colour.

Recognition false alarms are presented in column five of Table 2. As can be seen, responses to nonstudied items were noticeably lower than responses to studied items. False recognition was the same for the two instruction groups [$F(1, 68) = 2.71$, $MS_e = .021$, $p > .100$] and the two test form groups [$F(1, 68) = 1.33$, $MS_e = .021$, $p > .100$].

Remember Responses

A preliminary analysis involving the 32 conditions looked at the benefit of maintaining the study colour of tested pictures and words. This revealed one significant effect involving study/test colour: a marginally significant main effect [$F(1, 68) = 6.18$, $MS_e = .022$, $p < .050$], where items presented in the same colour at study and test (mean = .49) were remembered better than items presented in a different colour at study and test (mean = .45). All interactions involving colour failed to achieve significance, including the two-way interactions between colour and instructions [$F(1, 68) = 2.66$, $MS_e = .022$], test form [$F(1, 68) = .203$, $MS_e = .022$], study form [$F(1, 68) = 2.32$, $MS_e = .023$], or question context [$F(1, 68) = .853$, $MS_e = .019$] (all $ps > .1$). Other interactions involving colour that failed to achieve significance included: a three-way interaction between instructions and study form [$F(1, 68) = 1.56$, $MS_e = .015$], a three-way interaction between instructions and question context [$F(1, 68) = 1.80$, $MS_e = .019$], and a four-way interaction involving instructions, test form, and question context [$F(1, 68) = 2.14$, $MS_e = .019$] (all $ps > .100$). All other (three-, four-, and five-way) interactions involving colour were not significant (all Fs < 1). Thus, as with recognition, responses in "same" and "different" colour conditions were collapsed to simplify interpretation.

Remember hits – collapsed over colour.

Table 3 displays the mean proportion of correct remember responses, collapsed over colour, for the 16 conditions (four between- X four within-subject conditions). Inspection of the means in Table 3 revealed an overall picture over word advantage that, unlike with recognition responses, did not interact with the instruction groups (see rows five and six). Analysis confirmed this impression. There were significant main effects of instructions [$F(1, 68) = 4.63$, $MS_e = .162$, $p < .050$] and study form [$F(1, 68) = 79.23$, $MS_e = .037$, $p < .001$], but no significant interaction between them ($F < 1$). Thus, the picture over word advantage was the same for the no_questions group (difference = .22) and the questions group (difference = .20).

The three-way interaction between instructions, study form and question context was not significant, $F(1, 68) = .745$, $MS_e = .009$, $p > .1$. Planned comparisons (Fisher's LSD of .044) were nonetheless conducted as it was hypothesized that picture superiority would be reduced in the questions group only with congruent questions. However, contrary to the hypothesis, it was found that the advantage of study pictures over study words was significant in all conditions (differences of .20, .22, .20, and .19 for the congruent/no_questions, incongruent/no_questions, congruent/questions, and incongruent/questions conditions, respectively).

As observed with recognition responses, inspection of the means in Table 3 suggested that the advantage of congruent over incongruent context occurred only in the questions (row six), and not the no_questions (row five), group. Analyses confirmed that there was a significant main effect of question context [$F(1, 68) = 55.83$, $MS_e = .017$, $p < .001$], as well as an instruction X question context interaction [$F(1, 68) = 57.91$, $MS_e = .017$, $p < .001$]. Planned comparisons within groups (Fisher's LSD = .053) revealed that the benefit of congruent over incongruent context was significant in the questions group (difference = .228) but not the no_questions group (difference = .002).

There was a significant interaction between study form and test form [$F(1, 68) = 6.87$, $MS_e = .037$, $p < .050$], but, unlike for recognition responses, the two-way interaction between study form and question context was not significant ($F < 1$), as was the three-way interaction between study form, test form, and question context [$F(1, 68) = 1.70$, $MS_e = .009$, $p > .100$]. Post hoc comparisons (Fisher's LSD = .044) of the study form X test form interaction showed that the advantage of study pictures over study words was larger for items tested as pictures (difference = .26) than words (difference = .14).

As in recognition, the main effect of test form [$F(1, 68) = 1.01$, $MS_e = .048$] was not significant ($F < 1$), nor were its interactions with instructions ($F < 1$) and question context [$F(1, 68) = 1.53$, $MS_e = .017$, $p > .100$]. All other (two-, three-, and four-way) interactions were not significant (all Fs < 1).

In summary, remember responses showed an overall picture over word advantage that, unlike for recognition and contrary to our hypothesis, did not interact with instructions and/or question context.

Remember false alarms – collapsed over colour.

False alarms for remember responses are presented in column five of Table 3. As can be seen, remember responses to nonstudied items were noticeably lower than responses to studied items. False remember responses were the same for the two instruction groups [$F(1, 68) = 2.79$, $MS_e = .006$, $p > .050$] and the two test form groups [$F(1, 68) = 1.48$, $MS_e = .006$, $p > .100$].

Know Responses

Observed Frequencies of Know

A preliminary analysis (of the 32 study conditions) looking at the benefit of maintaining the study colour of tested pictures and words revealed no significant main effects or interactions involving colour. The main effect of colour was [$F(1, 68) = 3.10$, $MS_e = .022$] and the two-way interactions involving colour were [$F(1, 68) = .225$, $MS_e = .022$; $F(1, 68) = .063$, $MS_e = .022$; $F(1, 68) = 1.78$, $MS_e = .022$; and $F(1, 68) = 3.19$, $MS_e = .013$, for instructions, test form, study form, and question context, respectively (all $ps > .050$)] for observed know responses. Other interactions involving colour that failed to achieve significance were: a three-way interaction between instructions and study form [$F(1, 68) = 3.84$, $MS_e = .022$], a three-way interaction between test form and question context [$F(1, 68) = 1.45$, $MS_e = .013$], and a four-way interaction between instructions, study form and question context [$F(1, 68) = 3.39$, $MS_e = .014$] (all $ps > .100$). All other (three-, four-, and five-way) interactions involving colour were not significant (all Fs < 1). The data for responses in "same" and "different" colour conditions were collapsed in the remaining observed know analyses.

Observed know hits – collapsed over colour.

Table 4 displays the mean proportion of correct observed know responses, collapsed over colour, for the 16 studied conditions (four between- X four within-subject conditions). Inspection of the means revealed a different pattern than that for recognition and remember responses: there appeared to be no differences due to study form in either instruction group (see rows five and six). Analysis confirmed this impression, with no significant main effects of instructions or study form (both Fs < 1). However, there was a significant three-way interaction between instructions, study form and question context [$F(1, 68) = 7.54$, $MS_e = .007$, $p < .100$], as well as significant two-way interactions between instructions and study form [$F(1, 68) = 4.12$, $MS_e = .035$, $p < .050$], study form and question context [$F(1, 68) = 6.79$, $MS_e = .007$, $p < .050$], and instructions and question context [$F(1, 68) = 5.36$, $MS_e = .013$, $p < .050$]. Planned comparisons of the three-way interaction (Fisher's LSD = .039) conducted to test the hypothesis that picture superiority would be reduced in the questions group only with congruent questions found the reverse pattern. There was a significant negative advantage of study picture over study word in the congruent questions condition (difference = -.120), while positive but not significant differences of study picture/word in the incongruent questions, congruent/no_questions, and incongruent/no questions conditions (differences of .013, .025 and .022, respectively).

Unlike recognition and remember results, the main effect of question context [$F(1, 68) = 1.20$, $MS_e = .013$, $p > .050$] was not significant, nor was the main effect of test form ($F < 1$) and its interactions with instructions [$F(1, 68) = 3.05$, $MS_e = .132$, $p > .050$], study form ($F < 1$), and question context ($F < 1$). Other interactions that failed to achieve significance were: a three-way interaction between test form, instructions and study form [$F(1, 68) = 2.23$, $MS_e = .035$, $p > .050$] and a three-way interaction between test form, study form, and question context [$F(1, 68) = 2.90$, $MS_e = .007$, $p > .050$]. All other (two-, three-, and four-way) interactions were not significant (all Fs < 1).

In summary, a reversal of the picture superiority effect was observed in the congruent questions condition, while observed know responses for pictures and words were observed to be similar in all other comparisons.

Observed know false alarms – collapsed over colour.

Observed know false alarms are presented in column five of Table 4. As can be seen, know responses to nonstudied items were lower than responses to studied items. False observed know responses were the same for the two instruction groups and the two test form groups (both Fs < 1).

Estimated Probability of Know

The anomalies in the observed know findings (e.g., an absence of an effect of instructions and question context, with no interaction between test form and study form), raise questions about how to interpret the results. If we equate observed know responses as a direct measure of a specific process of memory, then we are implicitly assuming that the relationship between remember and this know memory process is one of perfect negative ($r = -1.00$) dependence (i.e., it assumes that one could not remember and know at the same time). Of course, the relationship between the two may not be one of negative dependence. It may be positive or neutral or somewhere in between the extremes. It would be useful to examine the interpretation of the observed know responses in models other than negative dependence. If one were to assume independence ($r = 0.00$) between remember and know, it would be necessary to estimate the probability of know using the following formula:

$$\text{est. } p(\text{know}) = \text{observed know}/[1 - p(\text{remember})]. \quad (1)$$

That is, the "estimated probability of know" would be calculated by dividing the proportion of know responses by the proportion not remembered, for each subject and each condition.

In a preliminary analysis of the estimated know responses looking at the benefit of maintaining the study colour of tested pictures and words, the main effect ($F < 1$) and all interactions involving colour were not significant. All two-way interactions involving colour were not significant, $F < 1$, as were the three-way interactions involving colour and test form and question context [$F(1, 52) = 3.48$, $MS_e = .044$] and instructions and question context [$F(1, 52) = 1.12$, $MS_e = .044$]. Other interactions that failed to achieve significance were: a four-way interaction between instructions, test form and question context [$F(1, 52) = 1.36$, $MS_e = .044$], a four-way interaction between instructions, study form and question context [$F(1, 52) = 1.13$, $MS_e = .041$], a four-way interaction between test form, study form, and question context [$F(1, 52) = 1.60$, $MS_e = .044$], and the five-way interaction between instructions, test form, study form, and question context [$F(1, 52) = 1.22$, $MS_e = .041$] (all $ps > .100$). All other (three-, four-, and five-way) interactions involving colour were not significant (all $Fs < 1$). Thus, the data for responses in "same" and "different" colour conditions were collapsed in the remaining estimated probability of know analyses.

Estimated probability of know hits – collapsed over colour.

Table 5 displays the mean estimated probability of know responses, collapsed over colour, for the 16 conditions (four between- X four within-subject conditions). Inspection of the means in Table 5 revealed an overall advantage of pictures over words that, like recognition and unlike remember responses, appeared to interact with the instructions group (see rows five and six). Analysis confirmed this impression. There were significant main effects of study form [$F(1, 63) = 23.59$, $MS_e = .040$, $p < .001$] and instructions [$F(1, 63) = 4.001$, $MS_e = .169$, $p = .050$], as well as a significant interaction between these variables [$F(1, 63) = 11.59$, $MS_e = .040$, $p < .010$]. Planned comparisons within groups (Fisher's LSD = .101) found that the advantage of study pictures over study words was not significant in the questions (difference = .202), but was in the no_questions (difference = .047) group.

As with recognition and remember responses, the three-way interaction between instructions, study form and question context was not significant ($F < 1$). Again, planned comparisons (Fisher's LSD = .091) were conducted to test the hypothesis that picture superiority would be reduced in the questions group only with congruent questions. Supporting the hypothesis, the advantage of study pictures over study words was smallest and not significant in the congruent question condition (indeed, a slight word advantage was found--difference = -.007), and significant and larger in the incongruent question, congruent/no_questions, and incongruent/no_questions groups (differences of .158, .246 and .103 respectively).

The effect of question context was significant [$F(1, 63) = 9.43$, $MS_e = .035$, $p < .010$], as was its interaction with instructions [$F(1, 63) = 9.95$, $MS_e = .035$, $p < .050$]. Planned comparisons within groups (Fisher's LSD = .094) revealed that the benefit of congruent over incongruent context was significant in the questions group (difference = .151), but not the no_questions group (difference = .002) (since subjects in the no_questions group were presented with "******" rather than congruent or incongruent questions).

The interaction between study form and question context was also significant [$F(1, 63) = 4.22$, $MS_e = .033$, $p < .050$]. Post hoc comparisons within groups (Fisher's LSD = .091) suggest that there was a benefit of congruent over incongruent context for study words (difference = .124), but not study pictures (difference = .024).

Like remember and unlike recognition responses, there was a significant interaction between test and study form [$F(1, 63) = 13.49$, $MS_e = .040$, $p < .001$], but the three-way interaction between test form, study form, and question context was not significant [$F(1, 63) = 2.54$, $MS_e = .033$, $p > .100$]. Post hoc comparisons within groups (Fisher's LSD = .101) of the test form X study form interaction found that the advantage of study pictures over study words was significant when tested as pictures (difference = .215) but not when tested as words (difference = .032).

As with recognition and remember responses, the main effect of test form was not significant ($F < 1$), nor were its interactions with instructions [$F(1, 63) = 4.00$, $MS_e = .169$, $p > .1$] and question context ($F < 1$). Other interactions that failed to achieve significance were: three-way interactions between instructions, test form, and study form [$F(1, 63) = 2.54$, $MS_e = .040$], instructions, test form, and question context [$F(1, 63) =$

2.24, $MS_e = 035$], and test form, study form, and question context [$F(1, 63) = 2.54$, $MS_e = .033$] (all $ps > .1$). All other (two-, three-, and four-way) interactions were not significant (all Fs < 1).

In summary, when independence was assumed to be the relationship between remember and estimated know responses, the estimated probability of know responses were similar to those of recognition responses in that the advantage of study pictures over words was significant in all conditions except for the congruent question condition, supporting the hypothesis that picture superiority would be reduced under congruent meaning conditions.

Estimated probability of know false alarms – collapsed over colour.

Estimated probability of know false alarms responses to nonstudied items are presented in column five of Table 5. The estimated probability of a know response to nonstudied items were clearly lower than to studied items. The estimated responses did not differ as a function of instruction and test form ($F < 1$).

Source Identification Responses

A preliminary analysis of the 32 study conditions looking at study/test colour found only one significant effect: a marginally significant colour X instructions interaction [$F(1, 68) = 2.25$, $MS_e = .014$, $p < .050$]. Post hoc comparisons of difference scores between "same" and "different" coloured items revealed that in the no_questions group, performance was significantly better for same items (mean = .037), while there were no differences in the questions group (mean = -.008), $t(70) = 2.29$, $p < .05$. The main effect of colour was not significant [$F(1, 68) = 2.20$, $MS_e = .014$ $p > .100$], nor were the two-way interactions involving colour [$F(1, 68)$, = .499, $MS_e = .014$; $F(1, 68)$, = .386,

MS_e = .012; and $F(1, 68)$, = 1.16, MS_e = .020 with test form, study form, and question context, respectively (all *ps* > .05)]. Other interactions involving test colour that failed to achieve significance were: a three-way interaction between instructions and question context [$F(1, 68) = 1.30$, $MS_e = .017$], a three-way interaction between study form and question context [$F(1, 68) = 3.57$, $MS_e = .012$], a four-way interaction involving instructions, study form, and question context [$F(1, 68) = 1.85$, $MS_e = .012$], and a four-way interaction between test form, study form, and question context [$F(1, 68) = 3.29$, $MS_e = .012$] (all *ps* > .100). All other three-, four-, and five-way interactions involving colour were not significant (all *F*'s < 1). Thus, as before, the responses in "same" and "different" colour conditions were collapsed for the remaining source identification analyses.

Source memory hits – collapsed over colour.

Table 6 displays the mean proportion of correct source identification responses, collapsed over colour, for the 16 study conditions (four between- X four within-subject conditions). Inspection of the means in Table 6 revealed a pattern similar to that of remember responses--an overall picture over word advantage that did not interact with instruction groups (see rows five and six). Analyses confirmed that there were significant main effects of instructions [$F(1, 68) = 16.07$, $MS_e = .061$, $p < .001$] and study form [$F(1, 68) = 86.61$, $MS_e = .037$, $p < .001$], but no significant interaction between them ($F < 1$). Thus, a similar significant advantage of study pictures over study words was seen for both the no_questions (difference = .23) and questions (difference = .18) groups.

The three-way interaction between instructions, study form and question context was not significant ($F < 1$). As before, planned comparisons (Fisher's LSD of .042) were nonetheless conducted because it had been hypothesized that picture superiority would be reduced in the questions group only with congruent questions. Contrary to the hypothesis, it was found that the advantage of study pictures over study words was significant in all conditions (differences of .17, .21, .22, and .24 in the congruent questions, incongruent questions, congruent/no_questions, and incongruent/no_questions groups, respectively).

As found in all previous results, the advantage of congruent over incongruent context in source identification responses occurred only in the questions, and not in the no_questions, group. This was confirmed by a significant main effect of question context [$F(1, 68) = 41.49$, $MS_e = .012$, $p < .001$] and an instruction X question context interaction [$F(1, 68) = 29.30$, $MS_e = .012$, $p < .001$], where the difference between congruent over incongruent context was significant (Fisher's LSD = .051) in the questions (difference = .153) but not in the no_questions group (difference = .013).

There was a significant interaction between study form and test form [$F(1, 68) = 37.74$, $MS_e = .037$, $p < .001$] and a significant three-way interaction between study form, test form and question context [$F(1, 68) = 14.58$, $MS_e = .008$, $p > .1$], but the interaction between study form and question context was not significant [$F(1, 68) = 1.57$, $MS_e = .008$, $p > .100$]. Post hoc analyses of the three-way interaction (Fisher's LSD = .042) revealed that the advantage of study pictures over study words was larger for items tested as pictures--especially in the incongruent, rather than congruent, context condition (difference of .40, as compared to .30)--than for items tested as words, which showed no study picture/word differences (difference of .01, and .04 in the congruent and incongruent conditions, respectively).

The main effect of test form was not significant ($F < 1$), nor were its interactions with instructions [$F(1, 68) = 1.53$, $MS_e = .061$, $p > .100$] and question context ($F < 1$). All other (two-, three, and four-way) interactions were not significant (all Fs < 1).

In summary, the pattern of source identification responses was similar to remember responses. That is, there was an overall picture over word advantage that did not interact with instructions.

Source memory false alarms – collapsed over colour.

False alarms for picture and word source responses are presented in columns five and six of Table 6. As can be seen, responses to nonstudied items were noticeably lower than responses to studied items. However, while there was no significant effect of instructions [$F(1, 68) = 2.71$, $MS_e = .041$, $p > .100$] or test form [$F(1, 68) = 1.33$, $MS_e = .041$, $p > .100$], there was a significant main effect of study form [$F(1, 68) = 24.98$, $MS_e = .026$, $p < .001$], where subjects were more likely to falsely respond word (mean = .027) than picture (mean = .014). Because of this guessing bias, false responses to the studied items were corrected by subtracting nonstudied responses from study responses categories one and three for pictures and categories two and four for words. These old/new difference scores were then reanalyzed in a 2 X 2 X 2 X 2 mixed-factors MANOVA. However, because the pattern of significant and nonsignificant effects was identical to the analysis of uncorrected source identification for studied items, the analysis is not reported here.

Four-Alternative Forced Choice Test

A subsidiary purpose of the experiment was to evaluate if there was incidental memory of study colour for pictures and words using a four-alternative forced choice

(4AFC) test of colour recognition. In the 4AFC test, subjects selected one of four coloured line drawings. In half of the study and half of the nonstudy test trials, one of the 4AFC colours had been seen in the previous test of recognition (test primed), while in the remaining study and nonstudy trials none of the 4AFC choices had been seen in the previous test of recognition (nonprimed). Thus, studied and nonstudied pictures and words were analyzed in test primed and nonprimed 4AFC responses.

Table 7 displays the mean proportion of 4AFC responses in all 40 test conditions--four between-subject (two instructions X two test form) by ten within-subject (two study form X two question context X two colour, plus nonstudied items) conditions. Visual inspection of the means in Table 7 revealed that responses to primed items (rows 1 to 4) were greater than those to nonprimed items (rows 5 to 8) for both studied (columns 1 to 4) and nonstudied (column 5) items. This effect for nonstudied items was confirmed in a 2 X 2 X 2 mixed-factors MANOVA, with test priming (primed versus nonprimed) as the within-subject factor, and instructions (questions versus no_questions) and test form (pictures versus words) as between-subject factors. There was a significant main effect of test priming [$F(1, 68) = 8.91$, $MS_e = .011$, $p < .010$], where primed (mean = .122) were greater than nonprimed (mean = .101). Because the effect of test priming should influence responses to studied as well as nonstudied items, the proportion for studied items was adjusted by subtracting responses to nonstudied items in an attempt to control for test priming.

Table 8 displays the mean proportion of corrected (studied items minus nonstudied items) 4AFC colour recognition for the four between- (two instructions X two test form) and four within-subject (two study form X two question context) conditions.

Inspection of these also suggested an advantage of test priming, with primed (rows 1 to 4) greater than nonprimed responses (rows 5 to 8), which appeared to interact with test form. This impression was confirmed in a 2 X 2 X 2 X 2 X 2 mixed-factors MANOVA of studied items, with instructions (questions group versus no_questions group) and test form (picture versus word) as between-subject factors, and test priming (primed versus nonprimed), question context (congruent versus incongruent), and study form (picture versus word) as within-subject factors. There was a significant three-way interaction between test form, test priming, and study form [$F(1, 68) = 13.65$, $MS_e = .013$], as well as significant main effects of test priming [$F(1, 68) = 8.23$, $MS_e = .058$] and study form [$F(1, 68) = 25.84$, $MS_e = .026$], and a significant two-way interaction between test and study form [$F(1, 68) = 15.95$, $MS_e = .026$, $p < .001$] (all *p*s < .001). Post hoc comparisons within groups for the three-way interaction (Fisher's LSD = .053) revealed that the advantage of primed over nonprimed responses was larger for items presented in the same form at study and test (differences of .113 and .034 for study pictures and words tested a pictures and .010 and .072 for study pictures and words tested as words, where same study/test form are underlined).

There was a significant main effect of instructions [$F(1, 68) = 11.24$, $MS_e = .113$, $p < .010$], where 4AFC colour recognition was better when subjects received no_question (mean = .164) than question (mean = .070) instructions at study. That is, attention to meaning appeared to decrease incidental memory of colour.

None of the other effects were significant. The two-way interactions involving instructions with test form [$F(1, 68) = 1.31$, $MS_e = .113$, $p > .100$] and study form [$F(1, 68) = 2.76$, $MS_e = .026$, $p > .100$] failed to achieve significance, as did three-way interactions between test form, test priming, question context [$F(1, 68) = 1.63$, MS_e = .019], instructions, study form, and question context [$F(1, 68) = 3.76$, $MS_e = .024$], test form, study form and question context [$F(1, 68) = 2.12$, $MS_e = .051$], and a four-way interaction between test form, study form, encoding context, and test priming [$F(1, 68)$ = 1.02, $MS_e = .020$] (all ps > .050). All other (two-, three, four-, and five-way) interactions were not significant (Fs < 1).

To provide further support for the notion of interactive encoding in incidental episodic memory, we had intended to test the hypothesis that incidental memory for colour would be better for pictures and words studied under congruent (rather than incongruent) question context conditions for remember (rather than know) responses. However, because there were so few "know" responses in the remember/know recognition test, it was impossible to properly compare correct colour recognition for remember and know items. Thus, we were unable to test the hypothesis.

In summary, incidental colour recognition (as measured by the 4AFC test) was better for primed items, especially those presented in the same form at study and test. In addition, colour recognition was better in the no_questions group than the questions group. This result is different from Hayman and Cribbie's (1998) finding that memory for task relevant sensory features (i.e., picture colour) was higher following congruent meaning and congruent colour processing than incongruent meaning/congruent colour, congruent meaning/incongruent colour, or incongruent meaning/incongruent colour. Manipulating explicit attention only to meaning (but not colour) appears to reduce, rather than increase, memory for colour. These results suggest that the link between meaning elaboration and incidental colour requires explicit attention to bind them in episodic memory. Since these findings are secondary to the main purpose of the present study (to examine the effect of meaning elaboration on picture/word memory), these findings will not be discussed further.

Discussion

There were five noteworthy results in the experiment: 1) a main effect of study form (i.e., a picture superiority effect) in recognition, source identification, remember, and estimated know responses, but not in observed know responses; 2) an interaction between study and test form, which was similar in size for pictures and words, in recognition, source identification, remember, and estimated know responses, but not in observed know responses; 3) no main effect of test form in recognition, source identification, remember, estimated know, and observed know responses; 4) an interaction between study instructions and question context, where congruent questions benefited memory for both pictures and words, while incongruent questions had no such effect, in recognition, source identification, remember, and estimated know responses, but not in observed know responses; and 5) a reduction of the picture superiority effect in the congruent question condition in recognition and estimated know responses, but not in source identification, remember, and observed know responses. These findings will be discussed in turn for each dependent variable: recognition, source identification, remember, observed know, and estimated know responses.

Recognition

First, a main effect of study form was found, where an advantage of study picture over study word was observed in both instruction groups, and in both test form groups. This is consistent with the picture superiority effect (Paivio, 1971), including findings of an advantage of study pictures over study words on tests of recognition (e.g., Madigan, 1983; Mintzer & Snodgrass, 1999; Weldon & Coyote, 1998, etc., but see Durso & Johnson, 1980, and Emmerich & Ackerman, 1980).

Second, an interaction between study and test form was found in recognition responses. That is, recognition was better in both instruction groups when there was a match between form at study and test (pictures at study and test, or words at study and test) than when there was a mismatch (pictures at study and words at test, or words at study and pictures at test). This is similar to other findings that a match between form at study and test decreases responses latencies (e.g., Emmerich & Ackerman, 1979; Sternberg et al., 1995) and increases recognition (e.g., Mintzer & Snodgrass, 1999). This result is also consistent with encoding specificity (Tulving & Thomson, 1973) and transfer appropriate processing (Morris et al., 1977; Weldon et al., 1989) explanations of memory, which emphasize the importance of the similarity between encoding and retrieval processes for maximizing retrieval success. According to these theories, memory performance will be better when there is a match between cues at study and test. In the present experiment, there was a greater match between study and test cues when the form at study and test was the same (i.e., the picture-picture and word-word conditions) rather than different (i.e., the picture-word and word-picture conditions); thus, recognition was better in the picture-picture and word-word (rather than the picture-word and word-picture) conditions.

More importantly, the interaction between study and test form was similar in size for pictures and words, if not larger for words--the mean increase in recognition performance in the word-word (relative to the word-picture) condition was .105, as compared to .053 in the picture-picture (relative to the picture-word) condition. This finding is consistent with Emmerich and Ackerman (1979), who found a similar increase in recognition performance for picture-picture and word-word conditions relative to picture-word and word-picture conditions, respectively. This suggests that pictures do not provide more or "better" cues to benefit retrieval than words. However, these results are inconsistent with findings that pictures are more sensitive to changes in visual form than words (Kolers et al., 1985), and are not easily interpreted within either encoding specificity (Tulving & Thomson, 1973) or transfer appropriate processing (Morris et al., 1977; Weldon et al., 1989) frameworks. Specifically, study pictures are assumed to be recognized better than study words because pictures provide more or "better" information than words at encoding. Thus, both transfer appropriate processing and encoding specificity would appear to predict that a pictorial encoding advantage (i.e., main effect of pictures at study) would lead to a greater study by test interaction for pictures than words, because pictures at test should also provide more and "better" retrieval cues for pictures than words at test, than words at test provide retrieval cues for words.

Third, no main effect of test form was found in recognition responses. That is, in both the question and no_question groups, there was no overall difference in recognition performance using pictures or words at test. Thus, the superiority of pictures over words was limited to the study form. Again, such a result is not easily interpreted within either encoding specificity (Tulving & Thomson, 1973) or transfer appropriate processing (Morris et al., 1977; Weldon et al., 1989) accounts of memory processing. There is no a priori reason to assume that the advantage of pictures is true of study encoding only, while encoding at test leads to similar retrieval cue strength for pictures and words. Both transfer appropriate processing and encoding specificity appear to predict that, by providing a better study/test match for these "superior" cues, pictures at test should produce better recognition.

Fourth, an interaction between study instructions and question context, where congruent questions benefited both pictures and words, while incongruent questions had no effect, was found in recognition responses. That is, this interaction reflected the fact that 1) the congruent condition in the question group increased recognition of both pictures and words relative to the incongruent question group, while 2) the incongruent condition in the question group had no effect on recognition relative to the no_question group. The latter finding is inconsistent with arguments that incongruent meaning can disrupt meaning elaboration for pictures and not words (Intraub & Nicklos, 1985; Smith & Magee, 1980). Intraub and Nicklos (1985) argued that semantic processing is "automatically initiated" for pictures and not words. Based on this notion, it was hypothesized that, in the present experiment, incongruent meaning questions would disrupt "automatically initiated" meaning encoding (and thus decrease memory performance) for pictures and not words. The finding that the incongruent condition in the question group had no effect on recognition relative to the no_question group is inconsistent with this hypothesis. That is, contrary to predictions, there was no negative benefit (or cost) of encoding incongruent meaning for pictures or words. This finding, as well as the finding that congruent questions increased recognition of *both* pictures and words, are also in contrast with Hayman and Dew's (1997) observations that encoding congruent meaning facilitated recognition for "read" conditions only and had no effect for "generate" conditions, while encoding incongruent meaning reduced the advantage of the "generate" condition and had no effect on the "read." effects. It appears therefore, that different mechanisms underlie the picture superiority (Paivio, 1971) and generation (Slamecka & Graf, 1978) effects.

Fifth, relative to the no_questions group, there was a reduction in the size of the picture superiority effect with congruent questions in the question group for recognition responses. That is, the advantage of study picture over study word was smaller with congruent questions in the question than in the no_question group (picture/word differences of .082 and .236, respectively) and was smaller with congruent questions than incongruent questions in the question group (picture/word differences of .082 and .173, respectively). Thus, with recognition responses, there was support for the hypothesis that increased elaboration of meaning at study would facilitate memory for words more than for pictures, thus reducing the picture superiority effect. This finding is similar to Grady et al. (1998), who observed a nonsignificant reduction of picture superiority following similar levels of picture/word meaning elaboration. However, the small but significant picture superiority effect observed in the present experiment is inconsistent with past observations that picture superiority in recognition can be eliminated (or even reversed) by encouraging similar levels of picture/word meaning elaboration (e.g., Durso & Johnson, 1980; Emmerich & Ackerman, 1979). For example, Durso and Johnson (1980) reported a failure to find the picture superiority effect following presentations of "function" questions (i.e., "What is the object used for?"), and reported a reversal of the picture superiority effect (i.e., a word superiority effect) following "explicit imagery" questions (i.e., "Create an image"). Similarly, Emmerich and Ackerman (1979), using response latencies as their dependent variable, reported a failure to find the picture superiority effect following presentations of "conceptual" questions (e.g., "Does it unlock things" for a KEY). The inconsistency between the present and past studies may be due to some or all of the following: 1) compression of effects due to near ceiling levels of recognition performance in Durso and Johnson's (1980) and Emmerich and Ackerman's (1979) studies, 2) the absence of a control for study/test form interactions in Durso and Johnson's (1980) study, and 3) the use of nonstandardized picture/word stimuli in Durso and Johnson's (1980) and Emmerich and Ackerman's (1979) studies. These possibilities will be discussed in the following paragraphs.

As discussed in the introduction, one of the problems of past studies was near ceiling levels of recognition performance. For example, in Emmerich and Ackerman's (1979) study, recognition in the "conceptual" questions condition was near ceiling (91% and 93% for pictures and words, respectively). Because recognition in the control condition (where no orienting questions were presented) was also near ceiling for pictures (90% recognition, as compared to 82% for words), additional benefit to pictures due to meaning elaboration would have been more difficult to detect. This possibility may have existed in studies that did not have a control group. For example, Durso and Johnson (1980) manipulated study-orienting questions in all conditions and found that recognition was at ceiling in both the "function" (93% for both pictures and words) and "explicit imagery" (91% and 95% for pictures and words, respectively) conditions. However, because the study had no control group, they could not evaluate whether the addition of meaning elaboration had a similar effect for pictures and words, which was one of the questions investigated in the present experiment. More critically, they could not rule out the possibility that picture recognition would have been at ceiling without meaning elaboration.

It is also possible that ceiling effects can lead to a failure to observe the picture superiority effect. To decrease the chance that ceiling effects would "cloud" the results of the present study, a lengthy "filler task" and several "recency buffers" were included in the design to increase the interval between study and test and (hopefully) reduce overall recognition performance. These manipulations were somewhat successful. Specifically, in the no_question condition of the present experiment, recognition was at 83% and 60% for pictures and words, respectively, as compared to 90% and 82% in Emmerich and Ackerman's (1979) study. However, near ceiling levels of performance were observed for congruent questions in the questions group. Specifically, in the condition of the present experiment that is comparable to Durso and Johnson's (1980) "function" condition (congruent context at study, words at test, recognition as the dependent variable), performance was near ceiling (95% and 89% for pictures and words, respectively), and similar to the levels reported by Durso and Johnson (i.e., 93% for both pictures and words). Thus, because the high levels of recognition observed in the present experiment were similar to those reported by Durso and Johnson (1980) and Emmerich and Ackerman (1979), we cannot conclude that the elimination of the picture superiority effect observed in their studies was due simply to compression of effects due to ceiling levels of performance.

While it is apparent that ceiling effects could cause a reduction of the picture superiority effect, there are reasons to doubt that this occurred in the congruent questions condition of the present experiment. In an analysis where the top half of the subjects in recognition performance were removed from the instruction groups, there was still a significant interaction between study form and study instructions, $F(1, 43) = 5.98$, $MS_e = .02$, where the study advantage of pictures over words in the no_questions group (mean difference of .29) was reduced in the questions group (mean difference of .18). Yet in another analysis, which removed any subject who had a ceiling levels of recognition in any of the four study conditions (leaving 12 and 16 subjects in the no_questions, picture and word at test, groups, and 8 and 8 subjects in the questions, picture and word at test, groups), there was still a reliable interaction between study form and study instructions, $F(1, 40) = 5.66$, $MS_e = .02$, where the study advantage of pictures over words observed in the no_questions group was reduced in the questions group (mean differences of .25 and .15, respectively, as compared to .24 and .13 when all subjects were included). In short, removal of subjects who had high recognition rates did not remove the interaction between study form and study instructions. Thus, by inference, there is no convincing evidence that the reduction in the picture superiority effect observed in the congruent questions condition of the present experiment was a direct artifact of ceiling levels of performance in recognition.

A second possible reason that the picture superiority effect could appear to be eliminated (or even reversed) following meaning elaboration is the lack of a control for study/test form interactions with words. Such a result can occur if pictures and words are manipulated at study but not at test, with only words being used as retrieval stimuli at test. For example, Durso and Johnson (1980) tested recognition of studied pictures and words using only aurally-presented words, while Grady et al. (1998) used only visually-presented words at test. This is often done in an attempt to reduce recognition performance and help avoid ceiling effects (Grady et al., 1998). However, in this situation, pictures at study are being tested with a different form of test stimuli, while words are tested with the same form of test stimuli. Thus, picture superiority effects are pitted against the study/test form interaction, which could potentially mask picture superiority. For example, in the present study, picture superiority at study was larger when pictures, rather than words, were used at test (means of .89 and .83, respectively). In studies where only words at test are used therefore, we would expect the picture superiority effect to be reduced--with or without meaning elaboration--simply because the study/test form interaction (i.e., pictures and words at study, words only at test) could in part offset the advantage of pictures over words. Thus, the interpretation of results is ambiguous in studies that find a small and/or nonsignificant picture superiority effect following a manipulation of study form and (aurally- or visually- presented) words only (e.g., Durso & Johnson, 1980; Grady et al., 1998).

The above analysis may also apply to studies that show a reduction, elimination, or reversal of the picture superiority effect following meaning elaboration in recall. For example, studies have reported that "semantic processing," (D'Agostino et al., 1977) "referential" tasks, and "explicit imagery tasks" (Durso & Johnson, 1980) eliminated the picture superiority effect in recall. However, recall tests can also present a study/test form interaction that might counteract the picture superiority effect because it involves the self-generation of retrieval cues. Specifically, if one were inclined to self-generate verbal (rather than pictorial) cues, there would be a match between encoding and retrieval cues for words (rather than pictures) that would counteract picture superiority in a manner similar to the word-word condition of a recognition test.

A third reason--in conjunction with the preceding argument--that the picture superiority effect could appear to be eliminated (or reversed) is the type of stimuli used as pictures. The stimuli used in the present experiment were taken from Snodgrass and Vanderwart's (1980) standardized set of black and white line drawings. The pictures have been standardized on four variables of central relevance to memory and cognitive processing: name agreement, image agreement, familiarity, and visual complexity. Studies which have failed to find a picture superiority effect have used materials that were less thoroughly evaluated and normed (e.g., they used pictures that were hand drawn by the authors using a felt tip pen and not extensively judged for name or image agreement; Durso & Johnson, 1980; Emmerich & Ackerman, 1979). It is possible that pictures that are not consistently identified will lead to a smaller picture superiority effect. A smaller picture superiority effect would, in turn, be more easily masked or hidden by other effects, such as the study/test form interaction discussed earlier.

Source Identification

With one exception, source identification responses were affected by the same experimental manipulations as recognition responses. Thus, as in recognition, source identification responses showed: 1) a main effect of study form (i.e., a picture superiority effect), 2) an interaction between study and test form which was similar in size for pictures and words, 3) no main effect of test form, and 4) an interaction between study instructions and question context, where congruent questions benefited memory for both pictures and words, while incongruent questions had no effect in the question group. However, source identification responses differed from recognition responses in one important way: unlike in recognition, there was no significant reduction of the picture superiority effect in the congruent question condition. That is, the advantage of study picture over study word in source identification was similar with congruent questions in the question and the no_question groups (picture/word differences of .174 and .221, respectively), and was similar with congruent questions and incongruent questions in the question group (picture/word differences of .174 and .206, respectively). Thus, unlike recognition, source identification results were inconsistent with the hypothesis that increased elaboration of meaning at study would facilitate memory for words more than for pictures, thus reducing the picture superiority effect. These results are consistent with Durso and Johnson (1980), who also found an advantage of study pictures over study words (means of .88 and .74, respectively) with source identification following their six "referential" tasks.

Remember

With one exception, remember responses were affected by the same experimental manipulations as recognition, and were the same as source identification responses. Thus, remember results showed: 1) a main effect of study form (i.e., a picture superiority effect), 2) an interaction between study and test form which was similar in size for pictures and words, 3) no main effect of test form, and 4) an interaction between study instructions and question context, where congruent questions benefited memory for both pictures and words, while incongruent questions had no effect. In terms of the expected reduction of the picture superiority effect in the congruent question condition, remember results essentially mirrored the pattern observed in source identification, rather than recognition, results. Specifically, like source identification but unlike recognition responses, the advantage of study picture over study word was similar with congruent questions in the question and the no_question groups (picture/word differences of .201 and .200, respectively), and was similar with congruent questions and incongruent questions in the question group (picture/word differences of .201 and .186, respectively). Thus, unlike recognition and like source identification results, remember results were inconsistent with the hypothesis that increased elaboration of meaning at study would facilitate memory for words more than for pictures, thus reducing the picture superiority effect.

The finding that remember and source identification responses were more similar to each other than to recognition responses is consistent with theories that see remember and source identification responses as relatively pure measures of episodic memory (e.g., Conway & Dewhurst, 1994; Dewhurst & Conway, 1994; Dewhurst & Hitch, 1999; Rugg et al., 1998; Wilding & Rugg, 1996), while recognition responses are similar but also include familiarity information. That is, theorists have conceptualized remember responses as representing recollection (based on the conscious retrieval of a specific study episode) (e.g., Gardiner & Java, 1993; Rajaram, 1993; 1996, etc.). Because source memory is, by definition, "the ability to remember the context in which a particular piece of information has been learned", (Rybash, et al., 1995, p. 112), we would expect behavioural results in source and remember responses to be similar, which is what occurred in the present experiment.

In contrast to remember and source responses, which appear to be pure measures of recollective experience (i.e., episodic memory), recognition responses are often hypothesized to be a composite of recollection and familiarity (Dewhurst & Hitch, 1999; Gardiner & Java, 1993; Rajaram, 1993; 1996). The finding that there was a reduction of the picture superiority effect in the congruent question condition (thus supporting the hypothesis that increased elaboration of meaning would facilitate memory for words more than for pictures) for recognition responses and not for remember or source identification responses suggests that picture/word differences in familiarity are supported by our manipulation of meaning, while picture/word differences in recollection are not. Durso and Johnson's (1980) results may also reflect the influence of familiarity in recognition—they found that meaning elaboration eliminated the picture superiority effect with recognition responses, but not source identification responses. Therefore, their recognition judgements could have reflected familiarity more than retrieval of a specific episode.

Evidence that picture/word differences in familiarity may be modified by manipulation of meaning may be evaluated by examining know responses, such as those used in the present experiment. If there is a dissociation between know results and remember/source results (i.e., know results show a reduction [or elimination] of the picture superiority effect in the congruent condition), the notion that our manipulation of meaning affected familiarity, but not recollection, is supported.

Observed Know

The observed know responses showed a pattern completely different from recognition, source identification, and remember responses. There was only one significant effect--a three-way interaction between instructions, study form and question context, which reflected a reversal of the picture superiority effect with congruent questions in the question group. Specifically, unlike recognition, source identification, and remember responses, memory for study words (mean = .339) was better than study pictures (mean = .219) in the congruent question condition.

The finding that a main effect of study for pictures was observed in remember but not know responses partially replicate Rajaram's (1993) finding of opposite picture/word memory performance in remember and observed know responses (i.e., she found picture superiority in remember responses and word superiority in observed know responses) and is consistent with evidence suggesting a dissociation between remember/know phenomena (e.g., Gardiner, 1998; Gardiner & Parkin, 1990; Macken & Hampson, 1991, Rajaram, 1993; 1996).

The anomalies in the observed know results (e.g., an absence of an effect of instructions and question context, no interaction between test form and study form, etc), raise questions about how to interpret the observed know responses. Simply analyzing observed know responses assumes that the relationship between remember and know memory processes is one of perfect negative dependence (r = -1.00). This might be true if remember and know responses are nonoverlapping areas in the same distribution (Donaldson, 1996), but if remember and know responses are separate measures of memory (i.e., their relationship is one of independence, r = 0.00), then know responses must be adjusted for overlap with remember responses. This was done in the "estimated know" results that follow.

Estimated Probability of Know

In contrast to observed know responses, estimated know responses were affected by the same experimental manipulations as recognition, and with one exception, remember and source identification responses. That is, estimated know responses showed: 1) a main effect of study form (i.e., a picture superiority effect), 2) an interaction between study and test form which was similar in size for pictures and words, 3) no main effect of test form, and 4) an interaction between study instructions and question context, where congruent questions benefited memory for both pictures and words, while incongruent questions had no effect. In terms of the expected reduction of the picture superiority effect in the congruent question condition, estimated know results showed a similar pattern as recognition results, only the effect of meaning elaboration was stronger. That is, the advantage of study picture over study word was smaller (in fact, slightly reversed) with congruent questions in the question than in the no_question group (picture/word differences of -.013 and .202, respectively) and was smaller with congruent questions than incongruent questions in the question group (picture/word differences of -.013 and .083). Thus, like recognition responses, estimated (assuming independence) know responses showed support for the hypothesis that increased elaboration of meaning at study would facilitate memory for words more than for pictures.

Summary

In summary, an overall picture superiority effect was observed in all dependent variables (recognition, source identification, remember, estimated know responses) except observed know responses. The effect was reduced in the congruent questions condition for recognition and estimated know responses, but not for source identification and remember responses, suggesting that our manipulation of meaning affected picture/word differences in familiarity, but not recollection.

Implications for Models of Picture Superiority

The following paragraphs will discuss the findings in terms of their implications for other explanations of the picture superiority effect--the dual-coding model of Paivio

(1971, 1986), and the sensory-semantic model of Nelson et al. (1977; Nelson, 1979). As discussed in the introduction, Paivio's dual-coding model suggests two bases for the picture superiority effect: First, pictures are more likely to be coded in both the verbal and image code than words, and, second, the image code produces a stronger memory trace than the verbal code. Nelson et al.'s model attributes the picture superiority effect to the superior sensory code of pictures, which makes them more "visually distinctive" than words. Direct evidence for (or against) these theories are not provided by our results, since the purpose of the present experiment was not to test specific predictions based upon them. However, in the context of the theoretical debate regarding the basis for the picture superiority effect, one variable of interest in our study is the relative "form change costs" observed for stimuli studied in picture form and stimuli studied in word form (i.e., differences in performance between items studied and tested in the same and different form). Since both dual-coding and sensory-semantic theories assume modality-specific benefits for pictures (i.e., pictures have a unique perceptual advantage because of their ability to tap into the stronger image code [Paivio] or their superior sensory features [Nelson et al.]), these theories appear to predict a greater relative increase for picture-picture conditions than word-word conditions. Contrary to these predictions, we found no main effect of test form, and an interaction between study and test form that was similar in size for pictures and words; that is, the mean increase in recognition performance in the word-word (relative to the word-picture) condition was .105, as compared to .053 in the picture-picture (relative to the picture-word) condition. Thus, the present results appear to (indirectly at least) contradict the dual-code and sensory-semantic theories.

Limitations of the Present Study and Suggestions for Future Studies

While our finding that the picture superiority effect in recognition was reduced (but not eliminated) was replicated internally (i.e., similar results were obtained for both picture and word test form groups and for both "high" and "low" recognition rates), strong conclusions cannot be made until the findings are replicated elsewhere. Durso and Johnson's (1980) findings, that the picture superiority effect in recognition can be eliminated, were also replicated internally--they reported similar results using six different "referential" tasks and two different "imaginal" tasks, although all estimates of picture superiority failed to account the interaction between study/test form. They also warned that picture/word memory results may differ depending on the semantic tasks used, and argued that: "Problems...may arise if a researcher decides to use only one semantic orienting task or decides to consider a number of types of semantic processing under the generic heading of semantic" (p. 423). Recall that the orienting questions used in our experiment covered a lot of areas (e.g., object animacy, function, physical descriptions, etc.) and ranged from specific descriptions with obvious answers (e.g., "animal that barks?" for a dog) to those that were more ambiguous (e.g., "back and forth?" for a swing). The heterogeneity of information covered by our semantic task may have lessened the chance we consistently tapped the encoding of meaning relations that contribute to the picture superiority effect. Consequently, it can be argued that our "diluted" semantic task was only able to reduce, but not eliminate, the picture superiority effect. This argument seems unlikely, since experiments that have used a variety of specific semantic tasks have reported little variance in their effectiveness (e.g., Durso & Johnson, 1980). Furthermore, this argument fails to explain why picture superiority in remember and source identification is apparently not affected by any semantic tasks. Despite this, future studies could investigate whether the present pattern of results hold up when different orienting tasks, including "imaginal" tasks (i.e., tasks which emphasize sensory information), are used.

Conclusion

The present study, which 1) took measures to control ceiling effects and assessed recognition at both "high" and "low" recognition rates, 2) included a manipulation of form at both study and test, and 3) used standardized picture and word materials, suggests that when potential confounds are controlled, the picture superiority effect in recognition is ubiquitous and robust, even following similar levels of meaning elaboration. Past evidence for the hypothesis that meaning elaboration accounts for the picture superiority effect is largely indirect, such as demonstrations that pictures access meaning more readily than words (e.g., Potter & Faulconer, 1975; Smith & Magee, 1980), and findings that semantically processed pictures and words activate similar brain areas (Grady et al., 1998). Therefore, there is no direct evidence that differences in meaning elaboration underlie the picture superiority effect.

References

Aggleton, J.P., Hunt, P.R., & Rawlins, J.N.P. (1986). The effects of hippocampal lesions upon spatial and non-spatial tests of working memory. Behavior and Brain Research, 19(2), 133-146.

Bajo, M.T. (1988). Semantic facilitation with pictures and words. Journal of Experimental Psychology: Learning, Memory, & Cognition, 14(4), 579-589.

Begg, I., Snider, A., Foley, F., & Goddard, R. (1989). The generation effect is no artifact: Generating makes words distinctive. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15(5), 977-989.

Conway, M.A., & Dewhurst, S.A. (1994). Remembering, familiarity, and source monitoring. Quarterly Journal of Experimental Psychology, 48A, 125-140.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

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.

Cribbie, R.A. (1995). The Effect of Semantic Representations on Episodic Memory For Unrelated Sensory Information. Master's Thesis, Lakehead University.

D'Agostino, P.R., O'Neill, & Paivio, A. (1977). Memory for pictures and words as a function of level of processing: Depth or dual coding? Memory & Cognition, 5(2), 252-256.

Dewhurst, S.A., & Conway (1994). Pictures, images, and recollective experience. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1088-1098.

Dewhurst, S.A., & Hitch, G.J. (1999). Cognitive effort and recollective experience in recognition memory. Memory, 7(2), 129-149.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. Memory & Cognition, 24(4), 523-533.

Durso, F.T., & Johnson, M.K. (1979). Facilitation in naming and categorizing repeated pictures and words. Journal of Experimental Psychology: Human Learning and Memory, 5(5), 449-459.

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

Emmerich, H.J., & Ackerman, B.P. (1979). The effect of orienting activity on memory for pictures and words in children and adults. Journal of Experimental Child Psychology, 28, 499-515.

Eysenck, M.W., & Keane, M.T. (1990). Cognitive psychology: A student's handbook. Hillsdale, NJ: Erlbaum.

Gardiner, J.M. (1998a). Functional aspects of recollective experience. Memory & Cognition, 16, 309-313.

Gardiner, J.M. (1988b). Generation and priming effects in word fragment completion. Journal of Experimental Psychology: Learning, Memory, & Cognition, 14, 495-501.

Gardiner, J.M., & Java, R.I. (1991). Forgetting in recognition memory with and without recollective experience. Memory & Cognition, 19(6), 617-623.

Gardiner, J.M., & Java, R.I. (1993). Recognizing and remembering . In A. Collins, M.A. Conway, S.E. Gathercole, & P.E. Morris (Eds.), Theories of memory (pp. 168-188). Hillsdale, NJ: Erlbaum.

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

Geis, M.F., & Hall, D.M. (1976). Encoding and incidental memory in children. Journal of Experimental Child Psychology, 22, 58-66.

Grady, C.L., McIntosh, A.R., Rajah, M.N., & Craik, F.I.M. (1998). Neural correlates of the episodic encoding of pictures and words. Proceedings of the National Academy of Sciences of the United States of America, 95, 2703-2708.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. Journal of Verbal Learning and Verbal Behavior, 23, 553-568.

Greene, R.L. (1986). Word stems as cues in recall and completion tasks. Quarterly Journal of Psychology, 38A, 663-673.

Haxby, J.V., Grady, C.L., Horwitz, B., Ungerleider, L.G., Mishkin, M., Carson, R.E., Herscovitch, P. Schapiro, M.B., & Rapoport, S.I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. Proceedings of the National Academy of Sciences of the United States of America, 88, 1621-1625,

Hayman, C.A.G., & Cribbie, R. (1998, June). Episodic memory for surface features of pictures is facilitated by study elaboration of meaning. Paper presented at the meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Ottawa Canada.

Hayman, C.A.G., & Dew, R.P. (1997, June). Facilitation from imagine versus perceive reflects greater attention to meaning. Paper presented at the meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Vancouver Canada.

Hayman, C.A.G., MacDonald, C.A., & Tulving, E. (1993). The role of repetition and associative interference in new semantic learning in amnesia: A case experiment. Journal of Cognitive Neuroscience, 5, 375-389.

Hayman, C.A.G., & Rickards, C. (1995). A dissociation in the effects of study modality on tests of implicit and explicit memory. Memory & Cognition, 23(1), 95-112.

Hayman, C.A.G., Servais, M., & MacDonald, C.M. (1997). Variations in the study processing of meaning affect episodic memory for surface features. Manuscript submitted for publication.

Hyde, T.S., & Jenkins, J.J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 12, 471-480.

Intraub, H. (1979). The role of implicit naming in pictorial encoding. Journal of Experimental Psychology: Human Learning and Memory, 5, 78-87.

Intraub, H., & Nicklos, S. (1985). Levels of processing and picture memory: The physical superiority effect. Journal of Experimental Psychology: Learning, Memory, & Cognition, 11(2), 284-298.

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

Job, R., Rumiati, R., & Lotto, L. (1992). The picture superiority effect in categorization: Visual or semantic? Journal of Experimental Psychology: Learning, Memory, & Cognition, 18(5), 1019-1028.

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

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

Kinjo, H., & Snodgrass, J.G. (2000). Is there a picture superiority effect in perceptual implicit tasks? European Journal of Cognitive Psychology, 12(2), 145-164.

Klein, K., & Saltz, E. (1976). Specifying the mechanisms in a levels of processing approach to memory. Journal of Experimental Psychology: Human Learning and Memory, 2, 671-679.

Kobayashi, S. (1986). Theoretical issues concerning superiority of pictures over words and sentences in memory. Perceptual and Motor Skills, 63, 783-792.

Kolers, P.A., Duchnicky, R.L., & Sundstroem, G. (1985). Size in the visual processing of faces and words. Journal of Experimental Psychology: Human Perception and Performance, 11, 726-751.

Macken, W., & Hampson, P. (1993). Integration, elaboration, and recollective experience. Irish Journal of Psychology, 14(2), 270-285.

Madigan, S. (1983). Picture memory. In J.C. Yuille (Ed.), Imagery, memory and cognition: Essays in honor of Allan Paivio (pp. 65-89). Hillsdale, NJ: Erlbaum.

Marks, W. (1991). Effects of encoding the perceptual features of pictures on memory. Journal of Experimental Psychology: Learning, Memory, & Cognition, 17, 566-577.

Menard, M.T., Kosslyn, S.M., Thomson, W.L., Alpert, N.M. & Rauch, S.L. (1996). Encoding of words and pictures: A positron emission tomography study. Neuropsychologia, 34, 185-194.

Metcalfe, J., Cottrell, G.W., & Mencl, W.E. (1992). Cognitive binding: A computational-modeling analysis of a distinction between implicit and explicit memory. Journal of Cognitive Neuroscience, 4, 289-298.

Mintzer, M.Z., & Snodgrass, J.G. (1999). The picture superiority effect: Support for the distinctive model. American Journal of Psychology, 112(1), 113-146.

Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. Nature, 273(5660), 297-298.

Morris, C.D., Bransford, J.D. & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Learning, 16, 519-533.

Moscovitch, M., & Craik, F.I.M. (1976). Depth of processing, retrieval cues, and the uniqueness of encoding as factors in recall. Journal of Verbal Learning and Verbal Behavior, 15, 447-458.

Nelson, D.L. (1979). Remembering pictures and words: Appearance, significance, and name. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory (pp. 45-76). Hillsdale, NJ: Erlbaum.

Nelson, D.L., Reed, V.S., & McEvoy, C.L. (1977a). Encoding strategy and sensory interference. Memory & Cognition, 5(4), 462-467.

Nelson, D.L., Reed, V.S., & McEvoy, C.L. (1977b). Learning to order pictures and words: A model of sensory and semantic encoding. Journal of Experimental Psychology: Human Learning and Memory, 3, 85-93.

Nelson, D.L., Reed, V.S., & Walling, J.R. (1976). Picture superiority effect. Journal of Experimental Psychology: Human Learning and Memory, 2, 523-528.

Nicholas, S. (1995). The picture-superiority effect in category-association tests. Psychological Research, 58, 218-224.

Nyberg, L. (1999). Imaging episodic memory: Implications for cognitive theories and phenomena. Memory, 7, 585-597.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). Mental representation: A dual coding approach. Hillsdale, NJ: Erlbaum.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255-287.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? Cognitive Psychology, 5, 176-206.

Parkin, A.J., Reid, T.K., & Russo, R. (1990). On the differential nature of implicit and explicit memory. Memory & Cognition, 18, 507-514.

Pellegrino, J.W., Rosinski, R.R., Chiesi, H.L., & Seigel, A. (1977). Picture-word differences in decision latency: An analysis of single and dual memory models. Memory & Cognition, 5, 383-396.

Peterson, S.E., Fox, P.T., Snyder, A.Z., & Raichle, M.E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. Science, 249(4972), 1041-1044.

Potter, M.C., & Faulconer, B.A. (1975). Time to understand pictures and words. Nature, 253, 437-438.

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

Rajaram, S. (1996). Perceptual effects of remembering: Recollective processes in picture recognition memory. Journal of Experimental Psychology: Learning, Memory, & Cognition, 22(2), 365-377.

Ritchey, G.H. (1980). Picture superiority in free recall: The effects of organization and elaboration. Journal of Experimental Child Psychology, 29, 460-474.

Rugg, M.D., Schloerscheidt, A.M., & Mark, R.E. (1998). An electrophysiological comparison of two indices of recollection. Journal of Memory and Language, 39, 47-69.

Rybash, J., Roodin, P. and Hoyer, W. (1995). Adult Development and Aging (Third Edition). Dubuque, IA: William C. Brown Publishers.

Scoville, W.B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery, and Psychiatry, 20, 11-21.

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

Slamecka, N.J., & Katsaiti, L. (1987). The generation effect as an artifact of selective displaced rehearsal. Journal of Memory and Language, 26, 589-607.

Smith, M.C., & Magee, L.E. (1980). Tracing the time course of picture-word processing. Journal of Experimental Psychology: General, 109(4), 373-392.

Snodgrass, J.G., & McLure, P. (1975). Storage and retrieval properties of dual codes for pictures and words in recognition memory. Journal of Experimental Psychology: Human Learning and Memory, 1, 521-529.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 6(2), 174-215.

Squire, L.R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychological Review, 99, 195-231.

Stein, B.S. (1978). Depth of processing reexamined: The effects of precision of encoding and test appropriateness. Journal of Verbal Learning and Verbal Behavior, 17, 165-174.

Sternberg, G., Radeborg, K., & Hedman, L.R. (1995). The picture superiority effect in a cross-modality recognition task. Memory & Cognition, 23, 425-441.

Sutherland, R.W., & McDonald, R.J. (1990). Hippocampus, amygdala, and memory deficits in rats. Behavior and Brain Research, 37, 57-79.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory (pp. 381-403). New York: Academic Press.

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

Tulving, E. (1985a). How many memory systems are there? American Psychologist, 40, 385-398.

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

Tulving, E. (1984). Multiple learning and memory systems. In K.M.J. Lagerspetz & P. Niemi. (Eds.). Psychology in the 1990's (pp. 163-184). Amsterdam, The Netherlands: Elsevier.

Tulving, E. (1993). What is episodic memory? Current Directions in Psychological Science, 2, 67-70.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Vaidya, C.J., & Gabrieli, J.D. (2000). Picture superiority in conceptual memory: Dissociative effects of encoding and retrieval tasks. Memory & Cognition, 28(7), 1165-1172.

Watkins, M.J., & Sechler, E.S. (1988). Generation effect with an incidental memorization procedure. Journal of Memory and Language, 27, 537-544.

Weldon, M.S., & Coyote, K.C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. Journal of Experimental Psychology: Learning, Memory, & Cognition, 22(3), 670-686.

Weldon, M.S., & Roediger, H.L., III (1987). Altering retrieval demands reverses the picture superiority effect. Memory & Cognition, 15, 269-280.

Weldon, M.S., Roediger, H.L., III, & Challis, B.H. (1989). The properties of retrieval cues constrain the picture superiority effect. Memory & Cognition, 17, 95-105.

Wilding, E.L., & Rugg, M.D. (1996). An event-related potential study of recognition memory with and without retrieval of source. Brain, 119, 889-905.

Wippich, W., Melzer, A., & Mecklenbrauker, S. (1998). Picture or word superiority effects in implicit memory: Levels of processing, attention and retrieval constraints. Swiss Journal of Psychology, 57(1), 33-46.

Zeki, S., Watson, J.D.G., Lueck, C.J., Friston, K.J., Kennard, C., & Frackowiak, R.S.J. (1991). A direct demonstration of functional specialization in human visual cortex. Journal of Neuroscience, 11, 641-649.

Appendix A

Remember/Know Test Instructions

You will be presented with pictures or words on the computer screen. As each picture or word appears, you will have to consider if you recognize it as having appeared in the study condition. If you recognize it you will then indicate, using the computer, whether or not you "REMEMBER" it from the previous list or just "KNOW" on some other basis that you saw it during the study condition. Additionally, you may indicate that you do not recognize the picture or word by responding "NOT studied as picture or word."

Each picture or word will be presented one at a time along with the following options:

1) Remember Picture from study session
2) Remember Word from study session
3) Know Picture (or similar one) was studied
4) Know Word (or similar one) was studied
5) Not studied as picture or word

* Note that the questions are asking you to indicate the form (picture or word) that you think the concept was presented in. For example, if you remember seeing a *picture* of a "tree" during the study session and now see the *word*, "tree," you would answer "1" or "3."

Please read the following instructions to clarify how to make "REMEMBER", "KNOW" and "NOT studied as picture or word" judgements.

Remember judgements: If your recognition of the picture or word is accompanied by a conscious recollection of its prior occurrence in the study manipulation, select a "REMEMBER" response (i.e., press (1) for "Remember Picture" or (2) for "Remember Word" on the keyboard). "REMEMBER" is the ability to become consciously aware of some aspects of the initial experience when the picture or word was previously presented (e.g.,, aspects of its physical appearance of the picture or word, a thought came to mind when you initially saw the picture or word, etc.).

Know judgements: If you recognize that the picture or word was presented during the study condition, but you cannot consciously recollect anything about its actual occurrence, select a "KNOW" response (i.e., press (3) for "Know Picture" or (4) for "Know word" on the keyboard). "KNOW" responses should be made when you are certain of recognizing the picture or word but do not have a specific conscious recollection of its occurrence in the study condition.

Not studied as picture or word: When you do not recognize the picture or word as appearing in the study list, you should indicate the number (5) for "Not studied as picture or word."

To further explain the difference between "remember" and "know" refer to these examples. If someone asks you what your name is, you would respond in the "know" sense without being consciously aware of anything about a particular event or experience. However, when asked what the name of the last movie you saw was, you would most likely respond in the "remember" sense. That is, you are consciously aware of some aspects of the previous experience. If you have any questions regarding these judgements feel free to ask the experimenter.

NOTE: THIS CAN BE A DIFFICULT TASK! WE DON'T EXPECT YOU TO GET EVERYTHING RIGHT. JUST DO THE BEST YOU CAN.

Thank You.

Footnotes

---

[1] Levels of processing proposes that focused attention to meaning information ("deep" levels of processing) results in better memory than reliance on sensory information. Transfer appropriate processing, on the other hand, explains memory performance as a function of the similarity between encoding and retrieval tasks. That is, focusing on the meaning (rather than the sensory) characteristics of an item will result in better memory performance for meaning tasks (and vice versa).

Table 1

Mean Proportion of Recognition Responses as a Function of Study and Test Form, Study Instructions, Question Context and Colour

Same Colour	Study Pictures		Study Words		Nonstudied
	Congruent	Incongruent	Congruent	Incongruent	
			No_Question[a]		
Test Pictures					
M	.833	.861	.539	.489	.147
SE	(.028)	(.042)	(.049)	(.067)	(.025)
Test Words					
M	.817	.772	.711	.583	.214
SE	(.045)	(.034)	(.035)	(.037)	(.047)
			Question		
Test Pictures					
M	.961	.861	.850	.600	.111
SE	(.014)	(.030)	(.036)	(.046)	(.041)
Test Words					
M	.939	.778	.906	.717	.147
SE	(.018)	(.041)	(.024)	(.045)	(.032)
Different Colour					
			No_Question[a]		
Test Pictures					
M	.850	.883	.550	.544	.161
SE	(.035)	(.035)	(.056)	(.054)	(.036)
Test Words					
M	.828	.794	.628	.711	.208
SE	(.040)	(.041)	(.050)	(.039)	(.046)
			Question		
Test Pictures					
M	.961	.872	.861	.561	.122
SE	(.018)	(.031)	(.037)	(.048)	(.031)
Test Words					
M	.961	.767	.878	.706	.127
SE	(.014)	(.040)	(.029)	(.045)	(.025)

Note. Recognition responses = remember and know responses summed.

[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 2

Mean Proportion of Recognition Responses as a Function of Study and Test Form, Study Instructions and Question Context

	Study Pictures		Study Words		Nonstudied
	Congruent	Incongruent	Congruent	Incongruent	
No_Question[a]					
Test Pictures					
M	.842	.872	.544	.517	.154
SE	(.026)	(.032)	(.049)	(.055)	(.028)
Test Words					
M	.822	.783	.669	.647	.211
SE	(.036)	(.031)	(.035)	(.031)	(.043)
Question					
Test Pictures					
M	.961	.865	.856	.581	.117
SE	(.012)	(.027)	(.034)	(.043)	(.034)
Test Words					
M	.950	.772	.892	.711	.138
SE	(.012)	(.033)	(.018)	(.042)	(.027)
Total					
No_Question					
M	.832	.828	.607	.582	.183
SE	(.022)	(.023)	(.032)	(.033)	(.026)
Question					
M	.956	.819	.874	.646	.127
SE	(.086)	(.022)	(.019)	(.031)	(.022)
Total					
M	.894	.824	.740	.614	.155
SE	(.014)	(.016)	(.024)	(.023)	(.017)

Note. Recognition responses = remember and know responses summed.

[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 3

Mean Proportion of Remember Responses as a Function of Study and Test Form, Study Instructions and Question Context

	Study Pictures		Study Words		Nonstudied
	Congruent	Incongruent	Congruent	Incongruent	
No_Question[a]					
Test Pictures					
M	.478	.536	.250	.264	.033
SE	(.065)	(.060)	(.035)	(.050)	(.011)
Test Words					
M	.558	.528	.386	.353	.065
SE	(.058)	(.055)	(.044)	(.051)	(.031)
Question					
Test Pictures					
M	.814	.600	.550	.314	.013
SE	(.051)	(.065)	(.071)	(.047)	(.068)
Test Words					
M	.658	.400	.519	.314	.025
SE	(.070)	(.046)	(.062)	(.052)	(.015)
Total					
No_Question					
M	.518	.532	.318	.308	.049
SE	(.043)	(.040)	(.030)	(.036)	(.016)
Question					
M	.736	.500	.535	.314	.019
SE	(.045)	(.043)	(.046)	(.035)	(.083)
Total					
M	.627	.516	.426	.311	.034
SE	(.033)	(.029)	(.030)	(.025)	(.092)

Note. Remember responses = remember picture and remember word responses summed.
[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 4

Mean Observed Frequency of Know Responses as a Function of Study and Test Form, Study Instructions and Question Context

	Study Pictures		Study Words		Nonstudied
	Congruent	Incongruent	Congruent	Incongruent	
		No_Question[a]			
Test Pictures					
M	.364	.336	.294	.253	.125
SE	(.057)	(.047)	(.043)	(.042)	(.028)
Test Words					
M	.264	.256	.283	.294	.156
SE	(.050)	(.042)	(.037)	(.038)	(.035)
		Question			
Test Pictures					
M	.147	.267	.306	.267	.107
SE	(.050)	(.064)	(.062)	(.037)	(.033)
Test Words					
M	.292	.372	.372	.397	.116
SE	(.074)	(.056)	(.056)	(.044)	(.023)
Total					
No_Question					
M	.314	.296	.289	.274	.141
SE	(.039)	(.032)	(.028)	(.028)	(.022)
Question					
M	.219	.319	.339	.332	.111
SE	(.046)	(.043)	(.042)	(.030)	(.020)
N					
Total					
M	.267	.308	.314	.303	.126
SE	(.030)	(.026)	(.025)	(.021)	(.015)

Note. Observed know responses = know picture and know word responses summed.
[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 5

Mean Estimated Probability of Know Responses as a Function of Study and Test Form, Study Instructions and Question Context

	Study Pictures		Study Words		Nonstudied
	Congruent	Incongruent	Congruent	Incongruent	
No_Question[a]					
Test Pictures					
M	.638	.772	.399	.350	.125
SE	(.061)	(.055)	(.056)	(.057)	(.028)
N	18	18	18	18	
Test Words					
M	.545	.503	.468	.434	.156
SE	(.064)	(.052)	(.055)	(.039)	(.035)
N	18	18	18	18	
Question					
Test Pictures					
M	.662	.586	.662	.385	.107
SE	(.110)	(.072)	(.078)	(.057)	(.033)
N	14	18	18	18	
Test Words					
M	.670	.585	.714	.581	.116
SE	(.084)	(.062)	(.063)	(.048)	(.023)
N	17	18	18	18	
Total					
No_Question					
M	.592	.638	.434	.392	.141
SE	(.044)	(.044)	(.039)	(.035)	(.022)
N	36	36	36	36	
Question					
M	.681	.586	.688	.483	.111
SE	(.065)	(.047)	(.050)	(.040)	(.020)
N	31	36	36	36	
Total					
M	.633	.612	.561	.437	.126
SE	(.040)	(.032)	(.035)	(.027)	(.015)
N	67	72	72	72	72

Note. Estimated probability of know responses = (frequency know/N) / (1-probability of remember).

[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 6

Mean Source Identification Responses as a Function of Study and Test Form, Study Instructions and Question Context

	Study Pictures		Study Words		Nonstudied	
	Congruent	Incongruent	Congruent	Incongruent	Pics[a]	Words[a]
Question[b]						
Test Pictures						
M	.783	.831	.469	.419	.078	.231
SE	(.029)	(.036)	(.048)	(.052)	(.022)	(.056)
Test Words						
M	.697	.644	.569	.572	.131	.292
SE	(.054)	(.051)	(.037)	(.040)	(.031)	(.068)
Question						
Test Pictures						
M	.925	.833	.647	.439	.100	.133
SE	(.016)	(.030)	(.052)	(.050)	(.040)	(.038)
Test Words						
M	.880	.700	.811	.683	.044	.231
SE	(.026)	(.037)	(.020)	(.040)	(.013)	(.049)
Total						
No_Question						
M	.740	.738	.519	.496	.104	.261
SE	(.031)	(.035)	(.031)	(.035)	(.019)	(.044)
Question						
M	.903	.767	.729	.561	.072	.182
SE	(.015)	(.026)	(.031)	(.038)	(.021)	(.032)
Total						
M	.822	.752	.624	.528	.088	.222
SE	(.020)	(.022)	(.025)	(.026)	(.014)	(.027)

Note. Source identification responses = remember picture and know picture responses summed for studied pictures and remember word and know word responses summed for studied words.

[a] The picture and word distinction for nonstudied items refers to the responses to these items, not the type of stimuli presented (i.e., pics = remember + know pictures; words = remember + know words).

[b] For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 7

Mean Proportion of Incidental Colour Recognition Responses as a Function of Study and Test Form, Instructions, Question Context and Colour

	Study Pictures		Study Words		Nonstudied
Primed	Congruent	Incongruent	Congruent	Incongruent	
			No_Question[a]		
Test Pictures					
M	.561	.617	.444	.400	.342
SE	(.052)	(.054)	(.038)	(.039)	(.023)
Test Words					
M	.472	.461	.450	.472	.231
SE	(.050)	(.038)	(.046)	(.047)	(.027)
			Question		
Test Pictures					
M	.517	.550	.372	.383	.336
SE	(.048)	(.040)	(.047)	(.036)	(.034)
Test Words					
M	.389	.322	.383	.406	.308
SE	(.051)	(.033)	(.043)	(.045)	(.024)
Nonprimed					
			No_Question[a]		
Test Pictures					
M	.417	.456	.378	.300	.281
SE	(.041)	(.045)	(.048)	(.027)	(.023)
Test Words					
M	.400	.450	.339	.328	.228
SE	(.050)	(.051)	(.031)	(.033)	(.031)
			Question		
Test Pictures					
M	.333	.306	.244	.261	.258
SE	(.044)	(.033)	(.025)	(.032)	(.020)
Test Words					
M	.306	.311	.283	.333	.242
SE	(.040)	(.033)	(.033)	(.039)	(.019)

Note. [a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).

Table 8

Mean Corrected (Studied – Nonstudied) Proportion of Incidental Colour Recognition Responses as a Function of Study and Test Form, Instructions, Question Context and Colour

	Study Pictures		Study Words	
Primed	Congruent	Incongruent	Congruent	Incongruent
No_Question[a]				
Test Pictures				
M	.219	.275	.103	.058
SE	(.043)	(.050)	(.034)	(.036)
Test Words				
M	.242	.231	.219	.242
SE	(.042)	(.043)	(.049)	(.045)
Question				
Test Pictures				
M	.181	.214	.036	.047
SE	(.044)	(.034)	(.051)	(.041)
Test Words				
M	.081	.014	.075	.097
SE	(.050)	(.045)	(.055)	(.047)
Nonprimed				
No_Question[a]				
Test Pictures				
M	.136	.175	.097	.019
SE	(.047)	(.050)	(.052)	(.039)
Test Words				
M	.172	.222	.111	.100
SE	(.059)	(.060)	(.049)	(.051)
Question				
Test Pictures				
M	.075	.047	-.014	.003
SE	(.046)	(.036)	(.022)	(.041)
Test Words				
M	.064	.069	.042	.092
SE	(.042)	(.036)	(.045)	(.033)

Note.
[a]For the no_question group, question context (congruent/incongruent) was a pseudo-manipulation (only "*******" was presented).