Available online at www.sciencedirect.com

SCIENCE DIRECT®

Consciousness and Cognition 11 (2002) 423–460

Consciousness and Cognition

www.academicpress.com

A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis

Dawn M. McBride[a,*] and Barbara Anne Dosher[b]

[a] *Department of Psychology, Illinois State University, Campus Box 4620, Normal, IL 61790-4620, USA*

[b] *Department of Cognitive Sciences and Institute of Mathematical Behavioral Science, University of California at Irvine, Irvine, CA 92697, USA*

Received 29 November 2001

Abstract

Four experiments were conducted to evaluate explanations of picture superiority effects previously found for several tasks. In a process dissociation procedure (Jacoby, 1991) with word stem completion, picture fragment completion, and category production tasks, conscious and automatic memory processes were compared for studied pictures and words with an independent retrieval model and a generate-source model. The predictions of a transfer appropriate processing account of picture superiority were tested and validated in "process pure" latent measures of conscious and unconscious, or automatic and source, memory processes. Results from both model fits verified that pictures had a conceptual (conscious/source) processing advantage over words for all tasks. The effects of perceptual (automatic/word generation) compatibility depended on task type, with pictorial tasks favoring pictures and linguistic tasks favoring words. Results show support for an explanation of the picture superiority effect that involves an interaction of encoding and retrieval processes.

* Corresponding author.
E-mail address: dmmcbri@ilstu.edu (D.M. McBride).

PII: S1053-8100(02)00007-7

1. Introduction

Increased memory performance for picture stimuli over word stimuli has been found in countless studies of recall and recognition. As a simple example of the effect, Paivio and Csapo (1973) had participants study pictures, concrete words, and abstract words under different encoding instructions (i.e., incidental or intentional). On a later free recall test, there was an effect of stimulus type such that pictures were better recalled than either type of word (concrete or abstract). Based on the numerous studies that have reported a picture superiority effect, there is no question that picture superiority exists under many conditions. What still is undetermined is the cause of the picture superiority effect. What aspect of pictures increases the likelihood that they will be recalled or recognized more often than words? Several theories have been proposed in an attempt to explain this phenomenon.

The current study was designed to test explanations of picture superiority. After studying picture and word stimuli, participants performed several memory tasks in a process dissociation procedure (Jacoby, 1991), allowing estimation of conscious and automatic memory processes and word generation and source matching processes by multinomial model fits. Comparisons of conscious and automatic memory parameters and word generation and source matching parameters for pictures and words tested predictions made by one class of theories of picture superiority.

2. Theories of picture superiority

2.1. Encoding theories

Two early theories suggested that picture superiority reflects a difference in the way that pictures are encoded. The first is the dual-coding theory, which was first proposed by Paivio (1975, 1986, 1991, 1995). Paivio claimed that picture stimuli held an advantage over words because they are dually encoded. While words are merely encoded verbally, pictures elicit both a verbal code and an image code because participants are more likely to generate a label for pictures than to image words. Having two types of codes connected to the pictures allows a greater chance of retrieval during a memory task. Given this assumption, dual-coding theory makes the prediction that the different codes, when activated, should have an incremental effect on performance.

The second encoding theory of picture superiority was the sensory-semantic theory suggested by Nelson and his colleagues (Nelson, 1979; Nelson, Reed, & McEvoy, 1977; Nelson, Reed, & Walling, 1976). According to Nelson (1979), pictures hold two encoding advantages over words. First, pictures are perceptually more distinct from one another than are words. Therefore, each picture is encoded more uniquely, increasing its chance for retrieval. Nelson et al. (1976) supported this idea with research investigating memory for pictures and words where the perceptual similarity of the pictures was manipulated. When similarity between items was low, pictures showed greater recall performance than words. When similarity was high, however, no picture superiority was evident. The second advantage, according to this theory, is that pictures access meaning more directly than words. If a semantic study instruction is given with study items (e.g., rate the pleasantness of each item), recall for words and pictures is similar (Paivio, 1975). In this case, words are encoded as "deeply" as pictures ordinarily are.

Although they differ somewhat in the advantage(s) proposed for pictures, the dual-coding and sensory-semantic theories both provide an explanation of picture superiority based on differences in encoding between pictures and words. Neither the dual-coding model nor the sensory-semantic model, however, directly predicts differences in picture or word memory due to differences in retrieval task. Because of the encoding emphasis of these theories, retrieval differences were not explicitly investigated in these early studies, and most of these studies used recall or recognition tasks to measure memory. These theories do not speak to comparisons between automatic and conscious memory.

2.2. *Transfer-appropriate processing*

More recently, Weldon and Roediger (1987) and Weldon, Roediger, and Challis (1989) discussed the picture superiority effect in the framework of transfer-appropriate processing (TAP). Transfer-appropriate processing theory states that a greater overlap of processing at study and test will result in improved performance on the test (Morris, Bransford, & Franks, 1977). For example, if items are encoded during a semantic task (i.e., a task that requires processing of meaning for the stimuli), performance should be higher for a memory test that relies on conceptual aspects of the items for retrieval than a test that relies on perceptual features. According to Roediger (1990), this is exactly why levels of processing effects are evident for recall and recognition tasks, but are absent for implicit tasks. Recall and recognition are explicit memory tasks that presumably require conceptual processing for retrieval, while typical implicit tasks may more often rely on perceptual processing (Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Schacter, 1987). If improved conceptual coding is the basis for the picture superiority effect, and implicit tasks do not rely on conceptual processing, then picture superiority should not be observed in implicit tasks.

Previous explanations of picture superiority implicated encoding as the significant process in the effect, but TAP accounts for picture superiority by an interaction of encoding and retrieval. If it is assumed that pictures are more likely than words to access meaning during encoding, performance for pictures should be higher on tasks that require conceptual retrieval. Weldon and Roediger (1987) claim that this is how superior performance for pictures on recall and recognition tasks can be explained. In addition, these authors show that for tasks that require perceptual processing of picture or word test items, performance reflects a match between the encoding and test item format: Studied words produce better performance on a word fragment completion task, while studied pictures are superior on tasks of picture fragment identification. These effects are presumably due to overlap of perceptual processing during study and test.

However, all implicit tasks do not show the same pattern of picture and word comparisons (see Table 1 for a summary of the results reviewed here). Weldon and

Table 1
Summary of the picture superiority results reviewed

Authors	Result	Task
Paivio and Csapo (1973)		
Concrete words	PS	Explicit recall
Abstract words	PS	Explicit recall
Weldon and Roediger (1987)	PS	Explicit free recall
Weldon et al. (1989)	PS	Explicit semantic cued recall
Weldon et al. (1989)	=	Explicit semantic fragment ID
Weldon et al. (1989)	WS	Explicit word fragment ID
Weldon and Coyote (1996)	PS	Explicit free recall
Weldon and Coyote (1996)	PS	Explicit category production
Weldon and Coyote (1996)	PS	Explicit word association
Wippich, Melzer, and Mecklenbrauker (1998)	PS	Explicit category production
Weldon and Roediger (1987)	PS	Implicit picture fragment ID
Weldon and Roediger (1987)	WS	Implicit word fragment ID
Weldon et al. (1989)	WS	Implicit word fragment ID
Weldon and Coyote (1996)	=	Implicit category production
Weldon and Coyote (1996)	=	Implicit word association
Wippich et al. (1998)	PS	Implicit category production
Wippich et al. (1998)	WS	Implicit word stem completion

Note. PS indicates picture superiority; WS indicates word superiority; = indicates no task performance difference for pictures and words.

Roediger's (1987) initial demonstration of word superiority over pictures on implicit tasks was followed by studies of different implicit and explicit tasks. Weldon et al. (1989) showed that explicit instructions were sufficient to produce a picture superiority effect: Picture recall was greater than word recall on a free recall test, but word recall was better than picture recall on both implicit and explicit word-fragment and word-stem completion tasks. In contrast, picture superiority was found for explicit cued-recall tasks where semantic word cues were given (i.e., a cue conceptually related to the target word). In addition, pictures and words showed equivalent performance on an explicit word-fragment completion task with semantically related words as cues, presumably because both perceptual and conceptual cues overlapped with the study episode. Although complex, these results were taken as support for a TAP approach to picture superiority.

Given the assumptions of TAP, pictures should show superior performance on implicit and explicit conceptual tasks. Weldon and Coyote (1996) recently used tasks believed to be conceptual in nature to compare memory for pictures and words. They confirmed a picture superiority effect for explicit free recall, category production, and word association tasks, but found equivalent performance for pictures and words on implicit category production and word association tasks. The results for the implicit tasks are in direct conflict with TAP predictions. Both implicit tasks showed levels of processing effects for pictures and words, which Weldon and Coyote presented as evidence that the tasks are indeed conceptual in nature, despite the unexpected lack of picture superiority for the tasks.

Wippich et al. (1998), on the other hand, found higher memory task performance for pictures than words on an implicit category production task with semantic study instructions. The effect was reduced with graphemic study instructions. Both the Weldon and Coyote (1996) and the Wippich et al. results rely directly on task performance. However, the use of a direct (conscious) retrieval strategy on the implicit tests cannot be ruled out, especially in the Wippich et al. study, where the implicit test results were affected to some degree by a level of processing manipulation.

In addition to these mixed results for conceptual implicit tasks, items encoded as pictures produce small amounts of priming (performance above baseline or unstudied levels) on word-based perceptual tasks such as word-fragment completion (Roediger, Weldon, Stadler, & Riegler, 1992; Weldon, 1993; Weldon & Roediger, 1987; Weldon et al., 1989). If implicit tasks such as word-fragment completion are based on perceptual processing, why would a study stimulus with no perceptual overlap to the test stimulus produce priming? Weldon and Jackson-Barrett (1993) showed that reducing fragment presentation time to 500 ms eliminated picture priming on a word-fragment completion task. This suggests that slow conceptual processes may contribute to fragment completion in some circumstances (see also Weldon, 1993).

Another possible explanation of picture priming on implicit perceptual tasks requiring a word response is that explicit processing contaminated the task (Bowers & Schacter, 1990). Results in implicit tasks that are inconsistent with TAP may reflect contamination by conscious processes. A recent procedure designed by Jacoby (1991) uses an elaborated procedure to segregate conscious and automatic memory rather than just measuring task performance in implicit and explicit tasks. The current experiments utilized Jacoby's procedure to estimate conscious and automatic memory for pictures and words and thereby test TAP explanations of the picture superiority effect and of theories that explain the effect as an interaction of encoding and retrieval processes.

In summary, explicit tasks such as recognition or recall yield picture superiority (see Table 1). However, the results for implicit memory tasks are not always consistent. Word superiority has been found for word-based perceptual implicit tasks (Weldon & Roediger, 1987; Weldon et al., 1989). However, Weldon and Coyote (1996) failed to find picture superiority for conceptual implicit tasks; whereas Wippich et al. (1998) found picture superiority for a conceptual implicit task (see Table 1). This study evaluates this inconsistency by using the process dissociation method to eliminate conscious contamination of implicit tasks.

3. Process dissociation procedure

In 1991, Jacoby proposed a process dissociation procedure (PDP) to estimate the amount of conscious and automatic processing that contributes to task performance. The procedure allows estimates that are said to be "process pure" based on performance in an inclusion task and an exclusion task. In the inclusion task, participants are instructed to produce either a studied response or any response, but are instructed on the exclusion task to never produce the studied response (i.e., they must produce an unstudied response). In this way, either type of processing can lead to the target response on the inclusion task, but conscious and automatic memory processes compete on the exclusion task. If conscious processes are successfully engaged during the exclusion task, the target response will not be produced; however, if automatic processes dominate, the target response will be produced.

Responding with a target item on the inclusion task is assumed to be due to consciously recollecting the item or having familiarity for the item (given that the item was not consciously recalled) due to automatic memory. The conscious and familiarity processes are assumed to be independent. Thus, the probability of a target response on the inclusion task is $C + (1 - C)A$. Following this logic, the production of a target item on the exclusion task can only occur if the item is not consciously recalled, but is brought to mind through automatic familiarity processes. Therefore, the probability of a target response on the exclusion task is $(1 - C)A$. From these equations, the probability of conscious recollection (C) and automatic familiarity (A) can be estimated from target production data in the inclusion and exclusion tasks.

The PDP has been used to argue that levels of processing manipulations differentially affect conscious and automatic memory processes (Toth, Reingold, & Jacoby, 1994)—that explicit memory is influenced by conceptual manipulations and implicit memory is influenced by perceptual manipulations (Schacter, 1987). Challis and Brodbeck (1992) have noted numerous studies where levels of processing effects were found for implicit tasks. Toth et al. (1994) argued that level of processing effects found for implicit tests are contrary to theoretical assumptions regarding the type of processing involved in conscious and automatic memory (see Schacter, 1987) and were due to contamination of explicit processing on the implicit tasks. Results from inclusion and exclusion tasks in the Toth et al. experiment indicated that the estimate of C was higher for the items studied under semantic instructions than nonsemantic instructions, but the estimates of A were equivalent for the two study conditions. These results supported the commonly held view regarding the type of processing underlying implicit and explicit tasks and the idea that a failure to find these dissociations for process impure implicit and explicit tasks may reflect some mixture of processing during the tasks. An analogous assertion for picture study items would be that picture superiority effects should not appear in process pure estimates of automatic memory for nonconceptual implicit tasks unless the task relied on perceptual processing for pictures.

The PDP relies on two major assumptions (Jacoby, 1991). The first is that the probabilities of consciously and automatically producing an item are equivalent for each task (inclusion or exclusion). The second is that conscious and automatic processing occur independently on any given trial. In other words, C and A are not correlated. If either of these assumptions is violated, inaccuracies in the estimation of C and A can occur.

The assumption of independence of C and A is currently debated (Cowan & Stadler, 1996; Curran & Hintzman, 1995; Jones, 1987 Joordens & Merikle, 1993). For example, Curran and Hintzman (1995) claim that under some conditions C and A are in fact correlated. In experiments utilizing the PDP, significant correlations between C and A estimates across participants and across items and patterns challenging parameter validity were found (see Jones, 1987; Joordens & Merikle, 1993, for plausible alternatives to independence, and Jacoby, 1998, for conditions of application). It is prudent to consider alternative accounts for the results gained from the process dissociation procedure, and in this article we consider an independent, direct-retrieval model based on the original analysis of Jacoby, as well as a generate-source model in which conscious recollection and automatic generation are dependent.

4. Direct-retrieval and generate-source models

Recently, multinomial process tree models (Batchelder & Riefer, 1990, 1999; Riefer & Batchelder, 1988) have been used as an extension and/or alternative to Jacoby's (1991) process dissociation model to estimate latent processes in a number of tasks (e.g., Buchner, Erdfelder, & Vaterrodt-Plünnecke, 1995). In these models, processing trees are developed that describe the possible processes that can lead to each particular response category. Different trees represent each task. Each branch in a tree is associated with a probability. Models are fit to response frequency data.

Multinomial models offer an attractive method of dealing with guessing parameters within the inclusion/exclusion task procedure, as parameters estimating guessing processes can be included. The process dissociation equations of Jacoby (1991), which assume independence, can easily be cast within a multinomial form. However, process and parameter dependence can also be incorporated within multinomial models.

Alternatives to the Jacoby independence model have been developed based on a generate-source conceptualization of various memory tasks that require the production of a word or item for response (e.g., Bodner, Masson, & Caldwell, 2000; Jacoby, 1998; McBride & Dosher, 1999; McBride, Dosher, & Gage, 2001). In the generate-source models, an item is first generated as a possible response and then the source of the generated item is evaluated (cf. Jacoby & Hollingshead, 1990).

In recent studies, researchers have described word-stem completion and cued-recall task performance with multinomial process tree versions of the generate-source model (Bodner et al., 2000; Jacoby, 1998; McBride & Dosher, 1999; McBride et al., 2001). Bodner et al. (2000) fit a generate-source model with four process parameters to response frequencies from a stem completion task. In their model, the probability of generating the target item using automatic memory, the probability of consciously recognizing a target given that it had been generated, and the probability of generating a target through guessing were all estimated. With these automatic and conscious parameters, independence is not assumed. This generate-source model provided good fits to data from three experiments with varying study instructions (e.g., generate, associate, and read) and different retrieval strategy instructions (i.e., direct-retrieval or generate-recognize).

Jacoby (1998) presented fits of a similar source model to stem completion data, where separate parameters were estimated for the probability of using an automatic process, a conscious process, and a guessing process to facilitate responses. Target and nontarget response frequencies were fit with a model that assumed that the automatic process occurred first with some probability (presumably comparable to an automatic response generation) after which the conscious process occurred with some probability (which can be seen as the source matching process in the generate-source model). Only under certain instructional conditions did the generate-source model fit the stem completion data well. McBride and Dosher (1999) found similar results in a comparison of word generation and source matching parameters for word stem completion and stem cued recall (this study is discussed further below). A comparison of the generate-source model with the original direct-retrieval model relies on qualitative comparisons for these nonnested models (e.g., Bodner et al., 2000; Jacoby, 1998).

5. The current study

The current experiments use the process dissociation to provide a test of an encoding–retrieval explanation of picture superiority as well as provide a further comparison of generate-source and direct-retrieval (PDP-based) models of production tasks. Encoding-retrieval theories (such as TAP) explain picture superiority as an interaction between encoding and retrieval processes for pictures and words, therefore, tasks requiring various amounts of conceptual and perceptual processing were given in a process dissociation procedure with picture and word stimuli. All tasks in the current study required production of a typed word in response to a specific cue. Conscious and automatic memory processes were estimated through the use of multinomial models. A generate-source model and a model modified from the original process dissociation equations (direct-retrieval) were separately fit to response frequency data in three experiments.

Although conceptual and perceptual processing are not exclusively linked with one form of memory or the other (e.g., Toth & Reingold, 1996),[1] researchers have argued that conscious memory is much more likely to be influenced by conceptual than perceptual processes (Toth et al., 1994). Likewise, it has been suggested that automatic forms of memory are more influenced by perceptual processing than conceptual processing (Schacter, 1987). There is a good deal of evidence to support these claims regarding conceptual and perceptual processing. For example, level of processing, a conceptual processing manipulation, affects conscious memory, but rarely affects automatic forms of memory. In addition, when level of processing has been shown to affect implicit task performance (e.g., Challis & Brodbeck, 1992; Weldon & Coyote, 1996), explicit contamination has not been ruled out. PDP studies that estimate conscious and automatic memory rather than rely on implicit and explicit task performance have generally shown little to no effect of conceptual manipulations on automatic memory estimates. These manipulations have included

[1] Studies have found some evidence of conceptual influences on automatic forms of memory (see Toth & Reingold, 1996, for a review). This point is discussed further under Section 10. Our argument here is not that conceptual–conscious and perceptual–automatic links are absolute. Instead, we argue that these links are stronger than the conceptual–automatic and perceptual–conscious links.

level of processing (Mecklenbraeuker, Wippich, & Mohrhusen, 1996; Toth et al., 1994) and attention (Schmitter-Edgecombe, 1999a,b).

Based on the links between conceptual and conscious processing and perceptual and automatic processing supported by these studies, the process estimates in the current study provided a test of interactive encoding–retrieval accounts of picture superiority. Specifically, predictions were made for the C and A estimates for items studied as pictures or as words.

5.1. The tasks

For the current experiments, a perceptual task, a conceptual task, and a task that relies on both types of processing were used. Using tasks that vary in the amount of conceptual and perceptual processing within the same general testing framework provides validity and consistency tests for the interactive encoding–retrieval account as an explanation of the pattern of picture superiority effects.

In Experiment 1, a free recall task was used to ensure that the stimuli in this study did in fact show picture superiority on a traditional explicit task.

In Experiment 2 of the current study, participants were given a picture fragment identification task, where participants must name fragmented pictures. Weldon and Roediger (1987) have shown that this task relies mainly on perceptual processing of pictures because picture fragments are given as test cues. Conceptual processing should contribute very little to this process.

In Experiment 3, participants were given a word-stem completion task, where they must produce a word that begins with three letters given by the experimenter, a task that is apparently more perceptual than conceptual. The greatest amount of priming has been shown for visually presented words (as compared to auditory words or pictures) on stem completion tasks, indicating that perceptual processing is prevalent, but a small amount of priming has been found for pictures and auditory words on these tasks, possibly indicating some conceptual influence (Craik, Moscovitch, & McDowd, 1994). In addition, word-stem completion has been shown to be similar to word-fragment completion in performance (Rajaram & Roediger, 1993; Roediger et al., 1992) and should therefore be representative of both types of verbal word tasks.

In Experiment 4, participants performed a category exemplar production task, in which exemplars of a category prompt are produced (Weldon & Coyote, 1996). Category production requires processing of the meaning of a generated item in order to determine if the item is a member of the given category. Past studies have shown that conceptual manipulations, such as levels of processing, affect performance on the category production task (Hamann, 1990; Srinivas & Roediger, 1990; Weldon & Coyote, 1996). For these reasons, it has been suggested that category production is a task that relies primarily on conceptual forms of processing for retrieval.

For all tasks, participants studied both pictures and words in an intentional study phase. At test participants received inclusion and exclusion instructions with a visual symbol (“O” or “N”) presented to indicate which task should be performed on each test trial. Different groups of participants received word-stem completion, category production, and picture fragment identification. For stem completion, participants were presented with a three-letter word stem. Inclusion instructions stated that the participant should try to produce a word they studied or the name of a picture they studied that began with the three letters given. If they could not produce a studied item, they were instructed to produce any word they could that began with the letter cue. In this case, if a studied item is given as a response, it could be because the participant consciously recalled it or automatically produced it. Exclusion instructions were to produce a word that began with the letters given that they had not studied either as a word or as a picture. They were instructed to be certain the word they produced was new to the experiment. For the exclusion task, if participants produced a studied word it should only be due to automatic memory because conscious recollection of study should not accompany any response.

The other tasks were framed in the process dissociation procedure in a similar manner. For picture fragment identification, fragmented pictures were presented at test and participants were asked to name them with a word or name of a picture from study or, failing recollection, any name they could think of for inclusion trials. For exclusion trials, they were asked to produce a name for the picture that was not a word or name of a picture presented at study. Picture fragments were piloted to ensure they were ambiguous enough to elicit more than one name, but not degraded to the extent that they could not be named. For category production, the name of a category and first letter of an exemplar were given and participants were asked to produce a studied exemplar or any exemplar for inclusion trials and a new exemplar for exclusion trials. For all exclusion trials, participants were asked to respond with "xxx" if they could not produce a word that they were certain was new to the experiment. They were also allowed to respond with "xxx" for inclusion trials if they could produce no response (studied or unstudied) to the test stimulus given. Therefore, there were three possible response categories for each task: targets, alternate words, and no responses (xxx).

5.2. TAP predictions

Based on prior evidence concerning the type of processing engaged during picture fragment completion, word stem completion, and category production, predictions for C (conscious) and A (automatic) estimates can be made according to TAP theory. For a task that relies primarily on conceptual processing, if it is assumed that pictures have superior conceptual coding compared to words (which is an assumption of TAP theory as discussed by Weldon & Roediger, 1987), the C estimates should be greater for picture than word stimuli due to overlap of the type of processing (conceptual) from study to test. The magnitude of the C estimates should be highest for the category production and stem completion tasks since these tasks are believed to require some amount of conceptual processing. Conscious estimates should be lower overall for picture identification due to the small amount of conceptual processing associated with this task.

According to claims of a strong perceptual influence, A estimates should depend on the amount and type of perceptual processing required for the task. For the stem completion task, which relies on partial word cues, the A estimate should be greater for words than pictures, since studied words provide a greater overlap of perceptual processing from study to test. Category production is primarily a conceptual task; therefore, the *A* estimates should reflect an advantage for studied pictures due to the conceptual processing advantage for picture items over word items. For the picture fragment identification task, which uses fragmented picture cues, the *A* estimate should be higher for pictures than words.

5.3. The models

Two multinomial models were fit to these data using the GPT program by Hu and Phillips (1999). Both model forms have been shown to fit data from word-stem completion tasks, but have been considered in past studies as particular models (Bodner et al., 2000; Jacoby, 1998; McBride & Dosher, 1999). A tree representation of each model can be seen in Figs. 1 and 2. One model (seen in Fig. 1) was a direct-retrieval model derived from the original process dissociation equations for stem completion (Jacoby et al., 1993; Toth et al., 1994) with word generation parameters

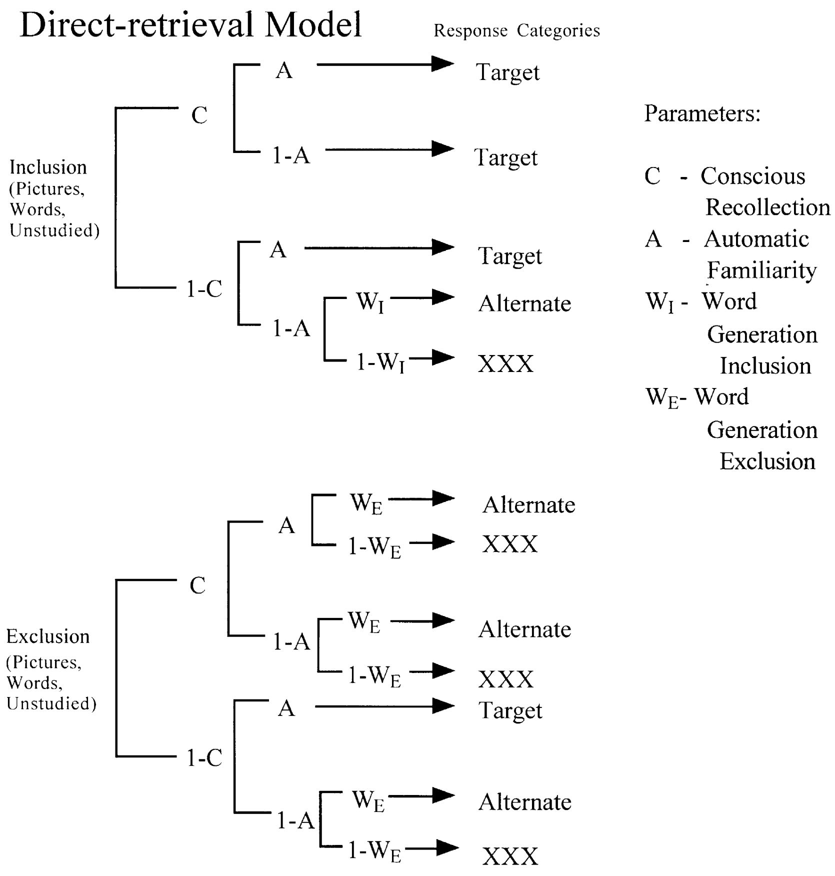

Fig. 1. Direct-retrieval processing tree model tested by McBride and Dosher (1999) based on Jacoby's (Jacoby, Toth, & Yonelinas, 1993) equations for production tasks.

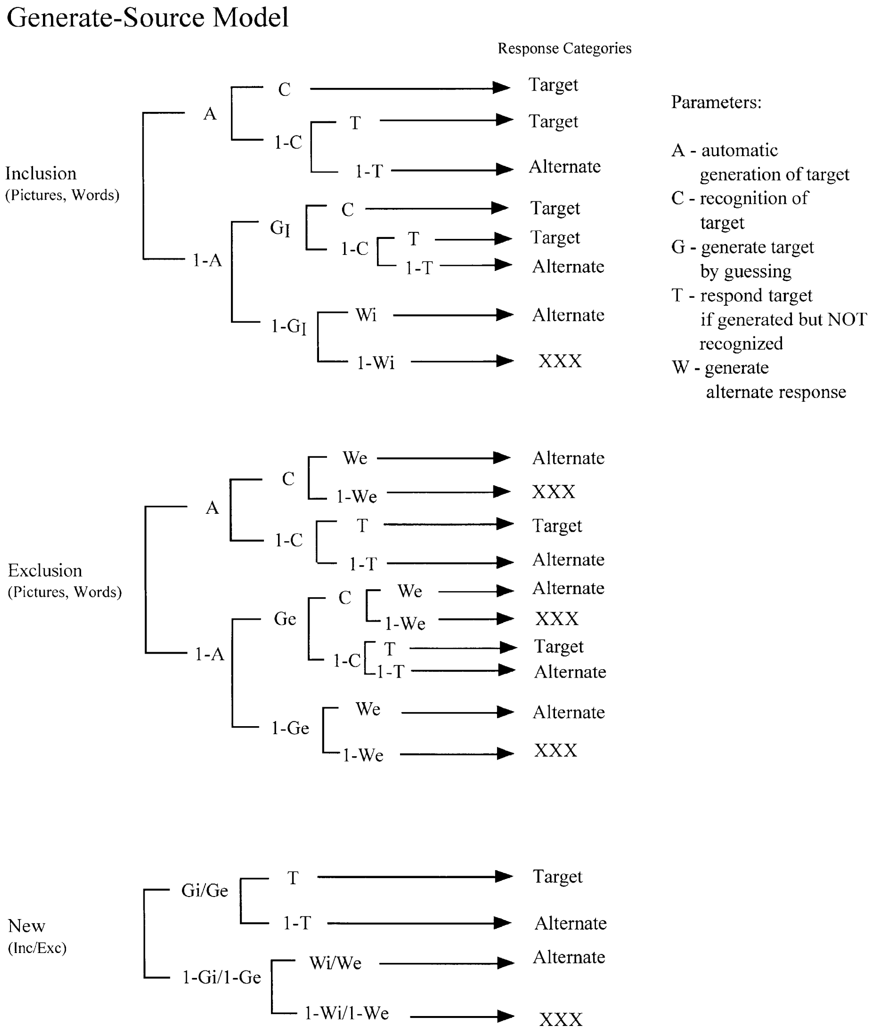

Fig. 2. Generate-source processing tree model based on Bodner et al.'s (2000) model.

added. Like the process dissociation equations, the direct-retrieval (Jacoby-based) multinomial model assumed independence between conscious and automatic processes on each trial.[2] Target responses were produced with the presence of conscious

[2] Buchner et al. (1995) (see also Buchner & Erdfelder, 1996) suggested that if one interprets the automatic memory estimate as a conditional probability, the independence assumption is not necessary. However, Wainwright and Reingold (1996) and Reingold and Wainwright (1996) state that this interpretation is just one specific form of the more general model suggested by the PDP equations. Therefore, an independence and a dependence model were each fit in the current study.

and/or automatic processing on the inclusion tasks and the absence of conscious processing on the exclusion tasks. This model is similar to the multinomial model tested by Jacoby (1998) and has been previously fit to stem completion and word fragment completion data by McBride and Dosher (1999) and McBride et al. (2001). A word generation parameter (W) was included in the model to estimate the probability of generating a nontarget word as a response on the tasks. The estimates for studied items must be compared with estimates of unstudied baseline values for interpretation (Jacoby et al., 1993).[3]

The second model (seen in Fig. 2) is based or a generate-source model of production tasks. Forms of this model were also previously tested by Jacoby (1998) and Bodner et al. (2000). McBride and Dosher (1999) also tested a related generate-source model. The generate-source model does not assume independence between conscious and automatic memory processes. This model assumes that participants automatically generate the target with some probability (A) and, if successful, attempt to determine the source of the item (e.g., studied or unstudied). If participants do not originally generate the target item ($1 - A$), they may attempt to guess the target item (G). If they either automatically generate or guess the target, they may or may not recognize it as a studied item (C). This model was extended to handle "xxx" responses used in the current experiments and to allow estimation of the probability of producing a nontarget item (W). For unstudied items, targets could only be given as responses through guessing. In other words, C and A were assumed to be 0 for unstudied items in this model, classifying the model as a high-threshold model. Bodner et al. (2000) added a parameter to this model to estimate the probability of responding with the target when it has been generated, but not recognized, as a studied item (T).[4] In the generate-source model, source matching (studied or unstudied) is assumed to be a conscious process (C), while target generation involves familiarity similar to the automatic process in the direct-retrieval model (A). Jacoby (1998) showed that a similar model to the one displayed in Fig. 2 fit stem completion data well when inclusion and exclusion tasks were given with generate-recognize instructions, while Bodner et al. found this model fit data well under varying instructional conditions. Both the direct-retrieval model and the generate-source model were fit to data from Experiments 2–4.

6. Experiment 1

Experiment 1 tested the picture superiority effect in a traditional explicit memory task to document the classic effect for the current stimulus set. Participants studied

[3] This comparison is consistent with a direct translation of the PDP equations into a multinomial model form and was the method used in the past studies employing this model (McBride & Dosher, 1999; McBride et al., 2001).

[4] As described by Bodner et al. (2000), this parameter was included to address the possibility that other possible completions (besides the target) are generated. The probability $1 - T$ indicates the likelihood that other completions come to mind and are chosen over the target.

80 items taken from a set of all stimuli used in Experiments 2–4 and were tested with a free recall task.

6.1. Method

6.1.1. Participants

Participants for Experiment 1 were 52 University of California, Irvine, undergraduate students who participated in none of the other experiments. Eleven additional participants took part in the norming of the picture stimuli. Participants volunteered for participation in exchange for course credit. All participants were native speakers of English.

6.1.2. Materials and design

The pictures for all experiments were taken from the Snodgrass and Vanderwart (1980) norms for picture stimuli. Words presented were single-word labels for pictures not shown. An original set of 120 pictures was normed with a separate group of participants ($N = 11$) to confirm common word labels. These participants were presented with pictures and asked to provide a one-word label for each picture. The set of pictures chosen had an average target word naming rate of 91.6% (i.e., on average participants produced the common label given in Snodgrass and Vanderwart 91.6% of the time), and no picture was used with an individual naming rate lower than 60%. See Appendix A for name agreement data by picture from both UC Irvine and Snodgrass and Vanderwart participants. A few items were later added to this set to accommodate the category production task (see Experiment 4). The complete target list contained 130 items.

For Experiment 1, the target list was composed of all items used in the later experiments ($N = 130$). For each participant, 80 items were chosen at random from the target set to be presented during the study phase. Forty items were randomly chosen to be presented as pictures, while 40 others were chosen to be presented as words. The 80-item study list ensured comparability with study list lengths used in the other experiments, providing similar on average study–test retention lags. Participants were run in individual cubicles, each containing a computer.

6.1.3. Procedure

Experiment 1 involved both a study and test portion. For the study portion, participants were presented with 80 randomly chosen items from the target list, 40 as pictures and 40 as words. Stimulus condition and presentation order for each item were determined randomly for each participant. Study items were each preceded by a fixation square. All stimuli were presented in white on a black background in the center of the screen. The cue fixation was presented for 1400 ms and each study item was presented for 3500 ms. Word items were shown in Times font with a 24-point font size. Pictures were all presized to fit into a 256×256 pixel-sized image with 75 pixels per inch. Participants were instructed to pay attention to each item and try to remember it.

During the test portion, participants were instructed to recall any items they could from the study portion (both words and pictures) in any order. Participants typed in one word for each item they recalled onto a blank screen. They were instructed to hit enter to go on to the next item or “xxxxx” when finished.

6.2. Results and discussion

An average of 16.96 pictures were correctly recalled ($SD = 5.67$), while an average of 12.52 words were correctly recalled ($SD = 6.47$). This difference was significant, $t(51) = 7.09$, $p < .001$, and indicates that these stimuli show the picture superiority effect on a traditional free recall task. Therefore, these stimuli are comparable to stimuli used in previous studies investigating picture superiority.

7. Experiment 2

Experiment 2 evaluated picture and word memory with a picture fragment identification task in a process dissociation procedure. Picture fragment identification as an implicit memory task has been shown to result in picture superiority (Weldon & Roediger, 1987). In this study, “process pure” conscious and automatic memory estimates were obtained through fits of the two multinomial process tree models described above (see Figs. 1 and 2).

7.1. Method

7.1.1. Participants

Fifty students voluntarily participated in Experiment 2 in exchange for course credit. All participants were UC Irvine undergraduate students and were native speakers of English. An additional 39 students participated in a pilot study to select the picture fragments used for this experiment.

7.1.2. Materials and design

Test items for Experiment 2 were fragmented pictures from the Snodgrass and Vanderwart (1980) picture norms. The fragments were chosen based on intermediate levels of target naming in a pilot experiment. Pictures were fragmented by randomly deleting 75% of the white pixels within the 256×256 pixel image. Thirty-nine participants viewed 130 fragmented pictures in different random orders and responded to each with a one-word name for the picture. The 130 normed pictures were the same items used in the other experiments in the current study. Participant responses were compared with labels given in Snodgrass and Vanderwart and 72 target items were chosen based on average naming rates and secondary labels given by the participants. Individual item completion rates for target items (with target labels) ranged from 5.1 to 61.5% with an average rate of 30.6%. This rate is comparable to average completion rates used in other picture fragment identification tasks (e.g., the fragments used by Weldon & Roediger, 1987, had an average baseline completion rate of 25%). The individual item rates are provided in Appendix B. These values indicate the percentage of participants who named the degraded item with its target label. For example, 33.3% of the pilot participants named the degraded image of a barn with its target label; 66.7% named the degraded barn with another label or could not name the item. In this case, lower values indicate a greater range of possible naming advantage due to memory. The pilot procedure ensured that fragmented pictures chosen for this task could be named with a consistent secondary label in the exclusion task.

7.1.3. Procedure

Participants were tested individually in small cubicles that contained personal computers. The experiment was divided into a study portion and a test portion. During the study portion, participants studied 48 of the target items selected during the pilot study. A different study set was randomly selected for each participant. Twenty-four of these items were presented as pictures and 24 were presented as words. In addition, participants studied 24 filler items (half as pictures and half as words) giving a total of 72 studied items to allow for similar average study–test delays as used in the other experiments. No filler items were tested. Each study item was preceded by a fixation square and was presented in the center of the screen. The cue fixation was presented for 1400 ms and each study item was presented for 3500 ms.

Participants received 72 test trials. Forty-eight trials were studied items (pictures and words) and 24 were unstudied items (not the same items used as fillers during study). For each trial, participants viewed a fragmented picture and responded with a one-word label according to test instructions. For inclusion trials, participants were instructed to name the fragmented picture with the name of a picture or with a word they saw in the first part of the experiment. If they were unable to remember a studied item that named the fragment, they were asked to name the fragment with any one-word label they could think of that represented what they saw or to type "xxx" if they could not think of a one-word label for the picture. For exclusion trials, participants were asked to name the fragments with a one-word label that was not presented (either as a word or a picture) during the study portion. They were to respond with a name for the fragment that was new to the experiment. A strategy was suggested to remember a word or picture presented earlier and then to come up with a different name for the picture as a response. If they could only think of a studied name or could not think of any name, they were asked to respond with "xxx." The pilot procedure confirmed possible alternative labels for the fragmented pictures. To indicate instruction, each picture fragment was preceded by an "Old" or "New" message for inclusion and exclusion, respectively. The instruction was displayed for 2870 ms and the fragment was displayed for a maximum of 28.5 s to allow participants enough time to begin a response. Responses were collected by computer keyboard.

For Experiments 2–4, participants received practice trials for the task. Before beginning the experiment, participants received four practice trials for both study and test portions. Participants studied four items, two as pictures and two as words. Then four test items were given, two with inclusion instructions and two with exclusion instructions. The practice items were items not shown during the experiment. None of the practice items could be used as responses for test items on any of the three tasks.

7.1.4. Response scoring

Each test response was compared with the corresponding target item. Responses that exactly matched the target item were computer-scored as matches. Responses that differed greatly (by more than three letters) from target items were computer-scored as mismatches. Typed responses that were fairly similar to target items (differing by three or fewer letters) were presented to a judge for scoring. The judge coded minor misspellings as matches. This scoring procedure was used for all experiments.

7.2. Results and discussion

Table 2 presents the observed frequencies for targets produced, alternate words produced, and no answer responses (i.e., "xxx") for both inclusion and exclusion

Table 2
Table of observed frequencies for Experiments 2–4

Condition	Target	Alternate	No response
Experiment 2—Picture fragment identification			
Inclusion			
Pictures	309 (0.515)	261 (0.435)	30 (0.050)
Words	191 (0.318)	370 (0.617)	39 (0.065)
Unstudied	123 (0.205)	423 (0.705)	54 (0.090)
Exclusion			
Pictures	98 (0.163)	423 (0.705)	79 (0.132)
Words	90 (0.150)	416 (0.693)	94 (0.157)
Unstudied	124 (0.207)	384 (0.640)	92 (0.153)
Experiment 3—Stem completion			
Inclusion			
Pictures	467 (0.531)	311 (0.353)	102 (0.116)
Words	545 (0.619)	252 (0.286)	83 (0.094)
Unstudied	314 (0.357)	403 (0.458)	163 (0.185)
Exclusion			
Pictures	64 (0.073)	771 (0.876)	45 (0.051)
Words	168 (0.191)	665 (0.756)	47 (0.053)
Unstudied	218 (0.248)	640 (0.727)	22 (0.025)
Experiment 4—Category production			
Inclusion			
Pictures	341 (0.573)	202 (0.339)	52 (0.087)
Words	299 (0.503)	233 (0.392)	63 (0.106)
Unstudied	204 (0.343)	289 (0.486)	102 (0.171)
Exclusion			
Pictures	81 (0.136)	408 (0.686)	106 (0.178)
Words	148 (0.249)	358 (0.602)	89 (0.150)
Unstudied	203 (0.341)	299 (0.503)	93 (0.156)

Note. Proportion values are presented in parentheses.

tasks for Experiments 2–4. In general, alternate responses given by the participants were concrete objects, indicating that the participants were attempting to complete the task according to the instructions. An ANOVA was conducted for the target data with test and study type variables. Main effects of test type (inclusion and exclusion), $F(1, 49) = 102.32$, $p < .001$, and study type (picture, word, and unstudied), $F(2, 98) = 45.16$, $p < .001$, were found to be significant. Inclusion trials ($M = .34$) resulted in more targets than exclusion trials ($M = .17$), a typical result for process dissociation procedure data. In addition, post hoc tests revealed that picture study ($M = .34$) resulted in higher overall target production than word study ($M = .24$). The two-way interaction of study and test type was also significant, $F(2, 98) = 34.32$, $p < .001$. Target production for unstudied items was equivalent on inclusion ($M = .205$) and exclusion task ($M = .207$), $t(49) = -0.062$, $p = .951$.

7.2.1. Model fits

The direct-retrieval and generate-source models displayed in Figs. 1 and 2, respectively, were fit to response frequency data (see Table 2). The resulting parameter estimates can be seen in Tables 3 and 4. For both model types, conscious and automatic estimates were higher for pictures than words. Nested models fits were performed to confirm the significance of the C and A differences. Nested models estimated either a single conscious or a single automatic component for both pictures and words. Tests were performed for both the direct-retrieval and generate-source models. G^2 comparisons for nested models for both multinomial model forms can be seen in Table 5.

Table 3
Estimates from the direct-retrieval model for Experiments 2–4

	Conscious	Automatic
Experiment 2—Picture fragments[a]		
Pictures	0.352 (0.025)	0.252 (0.019)
Words	0.168 (0.024)	0.180 (0.015)
Unstudied	0.001 (0.023)	0.206 (0.014)
Experiment 3—Stem completion[b]		
Pictures	0.458 (0.019)	0.134 (0.015)
Words	0.428 (0.021)	0.334 (0.018)
Unstudied	0.109 (0.021)	0.278 (0.013)
Experiment 4—Category production[c]		
Pictures	0.437 (0.025)	0.242 (0.021)
Words	0.254 (0.027)	0.333 (0.018)
Unstudied	0.002 (0.028)	0.342 (0.014)

Note. Standard deviations for the parameters calculated by the Hu and Phillips (1999) program are given in parentheses.
[a] $W_{inc} = 0.896$; $W_{exc} = 0.822$; $G(4)^2 = 3.16$.
[b] $W_{inc} = 0.735$; $W_{exc} = 0.948$; $G(4)^2 = 10.90$.
[c] $W_{inc} = 0.769$; $W_{exc} = 0.787$; $G(4)^2 = 5.49$.

Table 4
Estimates from the generate-source model for Experiments 2–4

	Conscious	Automatic
Experiment 2—Pictures fragments[a]		
Pictures	0.683 (0.160)	0.390 (0.044)
Words	0.531 (0.185)	0.143 (0.037)
Experiment 3—Stem completion[b]		
Pictures	0.839 (0.044)	0.270 (0.067)
Words	0.656 (0.062)	0.408 (0.030)
Experiment 4—Category production[c]		
Pictures	0.762 (0.049)	0.350 (0.044)
Words	0.504 (0.073)	0.243 (0.038)

Note. Standard deviations for the parameters calculated by the Hu and Phillips (1999) program are given in parentheses.

[a] $G_{inc} = 0.207$; $G_{exc} = 0.205$; $W_{inc} = 0.895$; $W_{exc} = 0.822$; $T = 0.9999$; $G(3)^2 = 3.15$.
[b] $G_{inc} = 0.357$; $G_{exc} = 0.248$; $W_{inc} = 0.735$; $W_{exc} = 0.948$; $T = 0.9999$; $G(3)^2 = 10.90$.
[c] $G_{inc} = 0.343$; $G_{exc} = 0.341$; $W_{inc} = 0.769$; $W_{exc} = 0.787$; $T = 0.9999$; $G(3)^2 = 5.49$.

Table 5
G^2 values for nested multinomial models

	Full model	$C_P = C_W$	$A_P = A_W$
Direct-retrieval model			
Experiment 2	3.16	30.24**	11.97**
Experiment 3	10.90	11.99	80.31**
Experiment 4	5.49	30.39**	16.09**
Generate-source model			
Experiment 2	3.15	9.60*	33.12**
Experiment 3	10.90	36.12**	25.07**
Experiment 4	5.49	35.99**	10.65*

* $G^2(1)$ difference is significant at $p < .05$.
** $G^2(1)$ difference is significant at $p < .01$.

Nested models comparing studied conditions significantly reduced the goodness of fit in all cases, therefore, all differences for Experiment 2 were significant, all ps $< .05$.

The direct-retrieval model A estimate for unstudied items (0.206) was similar to that for studied words (0.180), but lower than that for studied pictures (0.252), indicating positive priming for pictures and no priming for words. Nested models tests provided statistical confirmation of these results ($p > .05$ for words, $p < .05$ for pictures). Thus, the results show automatic priming for pictures on this task and insignificant automatic priming for words. In the generate-source model, no memory parameters (C or A) were estimated for unstudied items, therefore, priming is directly estimated by the value of A. A estimates were higher for pictures than words for this model. Both models indicated higher conscious estimates for pictures than words, as was expected. Using the AppleTree program by Rothkegel (1999), power analyses were conducted for each nested model comparison. Power was estimated for each nested model fit based on expected frequency data from the appropriate full model (i.e., direct-retrieval or generate-source model). For all power analyses, α was set at 0.005. Power was above 0.99 in all cases. Effect sizes ranged from 0.02 to 0.17.

Overall, Experiment 2 provided estimates that generally validated the past explicit and implicit memory task results (see Table 1) and extended these results to "process pure" or "latent" measures of automatic and conscious memory. The picture advantage was found in fits for both model types. This is consistent with claims that implicit or automatic memory are sensitive to the consistency between study and test formats (Schacter, 1987) in this picture fragment task. In addition, a task-independent conscious advantage for pictures was predicted from results for explicit tasks and supported by both the direct-retrieval and generate-source model fits. Experiments 3 and 4 explored stimulus format advantages for word-stem completion and category production memory tasks.

8. Experiment 3

A word-stem completion task was given in Experiment 3. This implicit task typically results in word superiority (Weldon et al., 1989). Conscious and automatic memory estimates were obtained through fits of the direct-retrieval and generate-source multinomial process tree models.

8.1. Method

8.1.1. Participants

Participants for Experiment 3 were 44 undergraduate students at UC Irvine who volunteered to participate in exchange for course credit. An additional four participants were run through the experiment; however, the data for these participants were not analyzed due to ceiling effects in the exclusion condition which can bias the *A* estimates (see Jacoby et al., 1993). All participants were native speakers of English.

8.1.2. Materials and design

The full set of normed target items was used as the stimuli for Experiment 3 (see Experiment 1 and Appendix A). Participants studied 80 pictures and words (40 of each type) randomly chosen from the full set ($N = 120$) that were presented in a random order on the computer screen. Stimulus condition and presentation order for each item were determined randomly for each participant. All items in the stimulus set began with a unique three-letter stem. Therefore, no two items shared the same stem. Participants received 120 test trials, 60 each for inclusion and exclusion tasks. For each task instruction (inclusion and exclusion), 20 items had been presented as pictures at study and 20 as words at study and 20 were new (unstudied). All stimuli were displayed on a personal computer.

8.1.3. Procedure

During the study portion, participants were presented with 80 picture and word items (in random order) and instructed to pay attention and try to remember each item. During the test portion, participants received 120 randomly ordered test trials. On each trial, a three-letter stem was presented, surrounded by two "O" symbols for inclusion trials or "N" symbols for exclusion trials. Each word stem was unique with respect to all items used in the experiment. Each stem had at least three possible word completions. Sixty trials were presented with inclusion instructions and 60 were presented with exclusion instructions. For inclusion trials, participants were instructed to complete the stem with a word they remembered from the study portion (either a word or the name of a picture presented) or if they could not remember a studied item to complete the stem with the first word they could think of. Again, a strategy of trying to remember a studied item to guide their response selection was suggested. If they could not think of any word that began with the letters presented, they were to respond with "xxx." For exclusion trials, participants were instructed to complete the stem with a word they knew was not presented during the study portion (either as a word or the name of a picture presented). If they remembered an item from the list, they were not to use it as a response. If they could not think of an unstudied item that began with the stem, the participants were instructed to respond with "xxx." For each trial, participants typed in a full word (not just the ending letters). Forty of the test trials represented items that had not been studied in the first portion of the experiment. The other 80 trials had been presented either as a word or a picture during the study portion (40 of each type).

8.2. Results and discussion

Participant responses were coded as in Experiment 2. Table 2 presents the observed frequencies for all conditions. Target data were analyzed in an ANOVA with study type (picture, word, or unstudied) and test instruction (inclusion and exclusion) variables. Main effects were found for both test, $F(1,43) = 306.67$, $p < .001$, and study type, $F(2,86) = 28.28$, $p < .001$. As was expected, target production for the inclusion test trials ($M = .50$) was higher than for exclusion test trials ($M = .17$). In addition, target production was higher overall for words ($M = .41$) than pictures ($M = .30$) as revealed by post hoc analysis. However, a significant interaction between test and study type was revealed, $F(2,86) = 74.86$, $p < .001$, indicating that although studied words produced more targets overall, this advantage differed for inclusion and exclusion trials. This finding is consistent with the differential predictions for conscious and automatic memory for the stem completion task. Target production was higher for unstudied items on the inclusion test ($M = .357$) than unstudied items on the exclusion test ($M = .249$), $t(43) = 3.83$, $p < .001$.

8.2.1. Model fits

Tables 3 and 4 present the process estimates for the direct-retrieval model and the generate-source model, respectively. For both models, the automatic estimates were higher for studied words than pictures. This result is consistent with word superiority effects found in previous studies of implicit word-stem completion (e.g., Weldon et al., 1989). The models differ, however, in the conscious estimates for the two stimulus types. For the generate-source model, the source matching parameter for target items was estimated to be higher for pictures than words (0.839 and 0.656, respectively) as predicted for this stem completion task by TAP. For the direct-retrieval model, the difference between C estimates for pictures and words was smaller (0.458 and 0.428, respectively) and nonsignificant. Despite this, both models show patterns of results that are consistent with past results of picture superiority for explicit memory.

Results from nested model comparisons for both multinomial models can be seen in Table 5. In all cases, power was estimated to be higher than 0.99. In addition, estimated effect sizes ranged from 0.01 to 0.17. Nested models significantly reduced the goodness of fit to the data in all fits (all ps $< .01$), except for the direct-retrieval $C_P = C_W$ model. In this case, a model that estimated one C parameter for pictures and words fit the data as well as the full model, indicating that the difference between conscious estimates for pictures and words was not significant. However, it should be noted that Weldon et al. (1989) found word superiority for explicit word-stem and word-fragment completion tasks. Therefore, it is possible that picture superiority for conscious memory is not as strong for tasks that involve fragmented words as retrieval cues.

With the current set of data for stem completion, the generate-source model displayed a larger difference between pictures and words for the estimate of conscious processing. Since the direct-retrieval and generate-source models are nonnested, they will only be compared qualitatively. The generate-source model estimates are most consistent with the TAP predictions for the stem completion task. In addition, target production was higher for unstudied items on the inclusion test than unstudied items on the exclusion test. Jacoby (1998) has argued that differences in baseline production between inclusion and exclusion tasks could indicate use of a generate-recognize strategy on the part of the participants. Both of these points suggest that the generate-recognize model may be the best one for the stem completion task.

In support of this view, there were several oddities in the estimates of the direct-retrieval model. The A estimate for items studied as pictures (0.134) was significantly lower than the A estimate for unstudied items (0.278), while the A estimate for items studied as words (0.334) was significantly higher. This result could indicate negative priming for the picture items. Last, the direct-retrieval model estimated the conscious memory estimate for unstudied items to be considerably greater than 0. Both of these findings raise questions about the external validity of the direct-retrieval model for stem completion performance.

9. Experiment 4

Experiment 4 evaluated picture and word memory in a process dissociation procedure with a category production task similar to the task used by Weldon and Coyote (1996) and Wippich et al. (1998). The goal of this experiment was to clarify mixed past findings for this task. Unlike the picture fragment identification and word stem completion tasks, this task presents conceptual retrieval cues at test. The contributions of conscious and automatic memory were estimated using multinomial models.

9.1. Method

9.1.1. Participants

Participants for Experiment 4 were 85 undergraduate students at the University of California, Irvine, who volunteered to participate in exchange for course credit. All participants were native English speakers.

9.1.2. Materials and design

Items for Experiment 4 were also chosen from the Snodgrass and Vanderwart (1980) picture norms. Items were chosen as exemplars of 14 different categories, with three exemplars representing each category. Within the categories, each target exemplar began with a unique letter. However, because test items for exclusion trials required participants to respond with an exemplar of the category that began with the letter given that had not been studied, for each target item, at least one other exemplar existed for that category that began with the same letter. All items used in Experiment 4 are given in Appendix C.

As in Experiments 2 and 3, participants received a study and test portion during the experiment. A new random assignment of items to conditions was derived for each participant. Within a category, one exemplar was assigned to each of the three study conditions: picture, word, and unstudied. Participants studied 28 target items (14 pictures and 14 words) and were tested on 42 items (the 28 studied items and 14 unstudied items). Each category was randomly assigned for each participant to either inclusion or exclusion instructions for the test portion, allowing seven categories for each instruction type. In other words, all three items associated with a single category (each representing a different study condition—picture, word, or unstudied) were tested with either inclusion or exclusion instructions for any one participant. Therefore, seven total items were randomly assigned to each of the six conditions in this experiment for every participant.

To allow for similar study–test retention intervals used in previous experiments, filler items not included in the target list were presented in random positions in the study list. Participants studied a total of 77 items, including the filler items. Half of the filler items were shown as pictures and half were shown as words. None of the filler items were tested. If a filler item was an exemplar of any of the target categories, it began with a unique first letter and therefore could not be used as a possible response for test trials. There were fewer test trials than in previous experiments, but these test trials took longer for participants to complete, again allowing similar average study–test delays.

9.1.3. Procedure

Participants for Experiment 4 were placed in individual cubicles containing a computer. The study portion contained 77 items and was similar to previous experiments. Each item was preceded by a fixation square and was presented as a picture (presized) or as a word. Timing and instructions for the study portion were identical to previous experiments.

Participants received 42 test trials. For each trial, a category name was presented along with the first letter of the target item and an "Old" or "New" signature above the category label to indicate inclusion or exclusion instruction, respectively. For inclusion trials, participants were instructed to respond with an exemplar of the category that began with the letter given that they had studied as a word or picture during the first part of the experiment. If they could not remember an item they had studied that fit the cues given, they were to respond with any appropriate exemplar. For exclusion trials, participants were to respond with an exemplar that fit the cues that they did not see (either as a picture or a word) in the study portion. They were instructed to respond only with items that were new to the experiment. As in previous experiments, participants were asked to first attempt to retrieve a studied item before responding. If participants could not think of a new item that fit the cues, they were to respond with "xxx." Test cues were displayed for a maximum of 28.5 s to allow participant to begin a response.

9.2. Results and discussion

Typed responses were scored in the same way as previous experiments. Response frequencies for each response category are shown in Table 2. In general, alternates given by the participants were exemplars of the category given that began with the correct letter, indicating that participants followed instructions for this task and were able to respond with appropriate targets. Target data were analyzed in a 3×2 ANOVA for study type and test instruction. A main effect was found for test type, $F(1, 84) = 143.94$, $p < .001$, such that target production was higher for inclusion ($M = .47$) than exclusion trials ($M = .24$). The main effect of study type was not found to be significant, $F(2, 168) = 1.42$, $p > .05$; however, study and test type did interact, $F(2, 168) = 51.46, p < .001$. Target production for unstudied items was equivalent on the inclusion ($M = .343$) and exclusion tasks ($M = .341$), $t(84) = .063$, $p = .950$.

9.2.1. Model fits

The two models seen in Figs. 1 and 2 were again fit to the response frequency data. Model estimates can be seen in Tables 3 and 4. For the direct-retrieval model, the C estimate for pictures was higher than that for words, indicating a conscious advantage for pictures. Similarly, the generate-source model indicated a picture superiority effect in conscious estimates for this task. The pattern of estimates of automatic memory was model-dependent. For the generate-source model, the A estimate for pictures (0.350) was higher than that for words (0.243), as expected under the TAP predictions for conceptually dependent category production. For the direct-retrieval model, the effect was reversed; the A estimate was higher for words (0.333) than pictures (0.242).

These results were tested statistically, as in previous experiments, by comparing nested model fits with the full model fits. All effects were significant, with all ps $< .025$. Power was greater than 0.99 for all nested models. Effect sizes ranged from 0.01 to 0.20. As in previous experiments, two comparisons for each model type (direct-retrieval and generate-source) were made to the full model, one nested model with a single conscious parameter for pictures and words and one nested model with a single automatic parameter for pictures and words. See Table 5 for the details of the model comparisons. For both model types, a model form with $C_P = C_W$ differed significantly from the full model indicating that pictures resulted in a higher C estimate than words. A model with $A_P = A_W$ also reduced the goodness of fit from the full model indicating that the A estimates differed for pictures and words (although in different directions for the two models). For the direct-retrieval model, the A estimate was higher for words, while the A estimate from the generate-source model was higher for pictures.

A qualitative comparison of the models suggests several differences. Unlike the data for Experiment 3, the baseline production rates for inclusion and exclusion are similar, which meets the requirements of the direct-retrieval model. Despite this, the inconsistencies in automatic estimates seen in Experiment 3 for the direct-retrieval model are also present here. For Experiment 4, the direct-retrieval model estimated the automatic parameter (A) for unstudied items (0.342) to be higher than the A estimates for pictures (0.242) and words (0.333). This difference was significant for the A_P vs A_U comparison (with a nested model equating these two parameters). These results could indicate negative priming for pictures and words, which is contrary to expected results. An alternate form of the generate-source model that does estimate automatic memory for unstudied items was fit to the data from Experiment 4 (see Section 10 for more details of this model). Automatic memory estimates for studied items were significantly above baseline (unstudied) estimates. Therefore, the generate-source model may be preferable here based on a qualitative comparison of the models.

Previous estimates from implicit tasks yielded mixed results—either a picture superiority effect or no effect for category production (Weldon & Coyote, 1996; Wippich et al., 1998). The current study evaluates these patterns using "process pure" or latent process estimates. Although the category production task in the current study used an initial letter cue, we believe that it is predominantly a conceptual task and hence should reflect picture superiority. Results from the generate-source model fits are consistent with this prediction. The generate-source model (both the original and alternate forms discussed above) estimated automatic memory to be higher for pictures than words, as is predicted for a conceptual task. Automatic memory estimates from the direct-retrieval model do not support this result, but are suspect due to negative priming results, which may indicate a validity problem for the model in the current study. On the other hand, the generate-source model estimates indicate picture superiority for conscious and automatic memory in the category production task.

10. General discussion

Overall, the current results are consistent with predictions based on an interactive encoding–retrieval framework such as the TAP model. Based on the link between conscious memory and conceptual factors described in Section 1, conscious estimates from the current model fits primarily indicate the influence of conceptual processing in each of the tasks for picture and word study items. In the three experiments using the process dissociation procedure, conscious memory estimates were higher for pictures than for words, indicating a conceptual advantage for pictures. This difference varied by task. Since the encoding phases of the three tasks were designed to be very similar, the variation in *C* estimates may be due to the greater use of conceptual processing in some tasks. In addition, picture superiority in estimates of conscious processing characterized both the direct-retrieval and generate-source models.

As described earlier, automatic memory estimates in picture fragment completion and in word stem completion have been shown to be largely influenced by perceptual manipulations. As expected, words showed higher automatic memory estimates than pictures for the stem completion task. This is predicted by TAP theory because test cues are fragmented words, which overlap more perceptually with studied words than studied pictures. In addition, pictures showed higher automatic memory estimates than words for the picture fragment identification task, a result also predicted by TAP theory due to the pictorial nature of the retrieval cue. The category production task is presumed to use little or no perceptual processing, therefore, automatic memory estimates should show the conceptual advantage for pictures. For the direct-retrieval model, automatic memory estimates favored words over pictures ($A_W > A_P$), while automatic memory estimates from fits of the generate-source model showed the predicted pattern ($A_P > A_W$). This is the only result where the model estimate patterns differed. Except for the *A* estimates in Experiment 4, both the direct-retrieval and generate-source models yielded similar results in each task.

Although the results of the direct-retrieval model are primarily consistent with models that highlight the role of the retrieval task in picture superiority, there are nonetheless significant irregularities in the parameter estimates that call the model into question. In some cases, the direct-retrieval model estimated memory process parameters to be lower for studied than unstudied items. One possible explanation for these results may be that participants used a generate-retrieval strategy on the tasks, making the direct-retrieval model inappropriate for these data. The generate-source model fit all of these data very well, yielded estimates that were almost uniformly consistent with the predictions of TAP, and posed no challenges to the external validity of the model. Although the higher target production rate in inclusion than exclusion tasks, sometimes associated with a generate-source strategy, was observed in Experiment 3, this is not a necessary signature of the strategy (Bodner et al., 2000).

Although the direct-retrieval model has some internal inconsistencies in estimated parameters, the automatic estimates are mostly consistent with characterizations in the literature of the explicit and implicit forms of the three tasks. The automatic memory (*A*) estimate for pictures was highest for the picture fragment identification task (a task that is believed to rely primarily on perceptual processing for pictures) and lowest for the stem completion task (a task believed to rely on perceptual word processing). The opposite was seen in the *A* estimates for word items. Stem completion resulted in the highest *A* estimate for words, while picture fragment identification gave the lowest automatic memory estimate for words.

In the generate-source model, as with the direct-retrieval model, the stem completion task resulted in higher source matching estimates (*C*) for both word and picture items than the other tasks. The picture fragment identification task gave lower source estimates for both picture and word items than the other tasks, which is expected if these estimates indicate the amount of conceptual processing required for this task relative to the other tasks used.[5]

The generate target estimates (A) for the source model were also as expected for word items. Stem completion showed the highest A estimates for word items, a result that is not surprising given that cues on this task contribute to fairly high word generation. The lowest A estimate for word items was for the picture fragment identification task. This task utilizes picture fragment as cues and is believed to rely on perceptual picture processing. Therefore, the picture fragment identification task showed a much higher A estimate for pictures than for words (0.390 vs 0.143, respectively). Predicted patterns were evident for the other tasks as well. The word-stem completion task resulted in a higher A estimate for words than pictures (0.408 vs 0.270, respectively), while the category production task showed the smallest difference in A estimates across the three tasks (0.350 for pictures and 0.243 for words). A estimates for pictures were fairly consistent across the three tasks, with predicted patterns supported for each task.

10.1. Comparison of the models

The multinomial models tested in the current study were based on models that have been tested in previous research. The direct-retrieval model was based on the original Jacoby equations for stem completion (see Jacoby et al., 1993) with the addition of parameters that estimated the probability of generating an alternate word. By estimating these W parameters, the model incorporated a form of word guessing. McBride and Dosher (1999) and McBride et al. (2001) evaluated this model in recent comparisons of forgetting in word-stem completion, word-fragment completion, and cued-recall tasks. Jacoby (1998) also tested a similar multinomial model for a stem completion task that allowed for guessing.

In the current study, the fits for this direct-retrieval model, however, showed some results inconsistent with theoretical expectations. First, the direct-retrieval model estimated the conscious unstudied parameter for the stem completion task to be considerably above 0. This result is unexpected since conscious recollection should not accompany unstudied items to this degree and cannot easily be explained. In addition, the direct-retrieval model fits resulted in paradoxical negative priming effects for the last two experiments. In Experiment 3, the automatic memory parameter for pictures was estimated to be lower than the automatic parameter for unstudied items. A similar result was also apparent in Experiment 2, where the automatic estimate for words was lower than that for unstudied items (nonsignificant). In Experiment 4, A estimates for both pictures and words were lower than the automatic

[5] The assumption that pictures have a conceptual advantage over words may seem to conflict with the classification of picture fragment completion as a perceptual task. However, this assumption is consistent with the TAP explanation of picture superiority proposed by Weldon and Roediger (1987). The estimate predictions made in this study follow from this assumption, but the automatic memory estimates found in these experiments would also be consistent with alternate forms of storage–retrieval theories that may not make such an assumption.

estimate for unstudied items (although the A_U vs A_W difference was not significant for Experiment 4). These results indicate negative priming where the automatic estimates for studied pictures or words are lower than the estimate for unstudied items. These findings have some precedent in the literature (e.g., Curran & Hintzman, 1995; Russo, Cullis, & Parkin, 1998) and have been taken as evidence of a violation of independence between C and A. Last, the direct-retrieval model provided estimates of conscious advantage for pictures that were not consistent across the three experiments, although the study conditions were approximately matched across experiments.

The relationships between the C_P and C_W estimated by the direct-retrieval model for the tasks are of interest. The ratios of the C_P to C_W parameter estimates for the last three experiments were 2.10, 1.07, and 1.72, respectively.[6] These values seem to be related to the expectation of picture superiority for the task. For example, in stem completion, word retrieval cues are given and word superiority is typically found for this task. In the current study, the C_P/C_W ratio for stem completion (for estimates from the direct-retrieval model) was lower than that for the other tasks. On the other hand, the picture identification task typically results in a strong picture superiority effect and also results in the highest C_P/C_W ratio. For direct-retrieval model fits, the type of retrieval cue given on the task appears to be related to the magnitude of the conscious memory advantage. This result is not consistent with findings indicating that conscious memory estimates are not sensitive to perceptual retrieval cues (Jacoby, 1996; Jacoby et al., 1993). The conscious advantage for pictures covaries with the automatic advantage for pictures in Experiments 2–4. Overall, these paradoxical results present a challenge for the direct-retrieval model.

The inconsistencies described here for the direct-retrieval model are related to those reported by other researchers. For example, Bodner et al. (2000) showed that in some cases A can be underestimated by the process dissociation equations, resulting in A estimates for studied items that are below baseline, even when participants are tested with direct-retrieval instructions. This result was also found by Curran and Hintzman (1995) for stem completion and Russo et al. (1998) for word-fragment completion. In addition, Bodner et al. claim that the recognition criterion used by the participants may be more important for completion rates on a stem completion task than the retrieval orientation they are instructed to use (e.g., direct retrieval or a generate-recognize strategy; see also Jacoby, 1998). In several experiments, they showed that C and A estimates from a stem completion task are consistent with a generate-recognize strategy even when participants are instructed to use a direct retrieval strategy. This result is consistent with the conclusion for the current experiment involving word-stem completion, where direct-retrieval instructions were given, but results from the direct-retrieval model were inconsistent with priming expectations.

On the other hand, the generate-source model fairs better on an examination of the estimate results. The generate-source model form discussed by Bodner et al. (2000) does not estimate C and A parameters for unstudied items (it is a high-threshold model); therefore, negative priming inconsistencies found for the direct-retrieval

[6] Ratios are useful in revealing the relative strength of conscious memory for pictures and words regardless of possibly large task variation in the level of importance overall given to conscious processing.

model are precluded by this high-threshold model. However, an alternate form of the generate-source model previously fit by McBride and Dosher (1999) and McBride et al. (2001) to data from word-stem and -fragment completion tasks does estimate memory parameters for unstudied items (this model can be seen in Fig. 3). Like the Bodner et al. model, this alternate model also assumes that retrieval on the tasks follows a generate-source strategy. Fits of the alternate generate-source model showed no negative priming for the automatic estimates (parameter estimates for this model are given in Appendix D for Experiments 2–4). Instead, studied picture and word items showed positive priming relative to unstudied items in all three tasks. Therefore, it appears that the direct-retrieval model may be unique in this inconsistent result. The two forms of the generate-source model (Bodner et al. form and McBride and Dosher alternate form) each show expected results. For Experiment 2, automatic memory estimates were lower for words than for pictures due to the use of picture fragments as test cues. In Experiment 3, the automatic estimate for words was higher than that for pictures. This result is expected since the stem completion task provides partial words as cues at test. In Experiment 4, automatic estimates were higher for

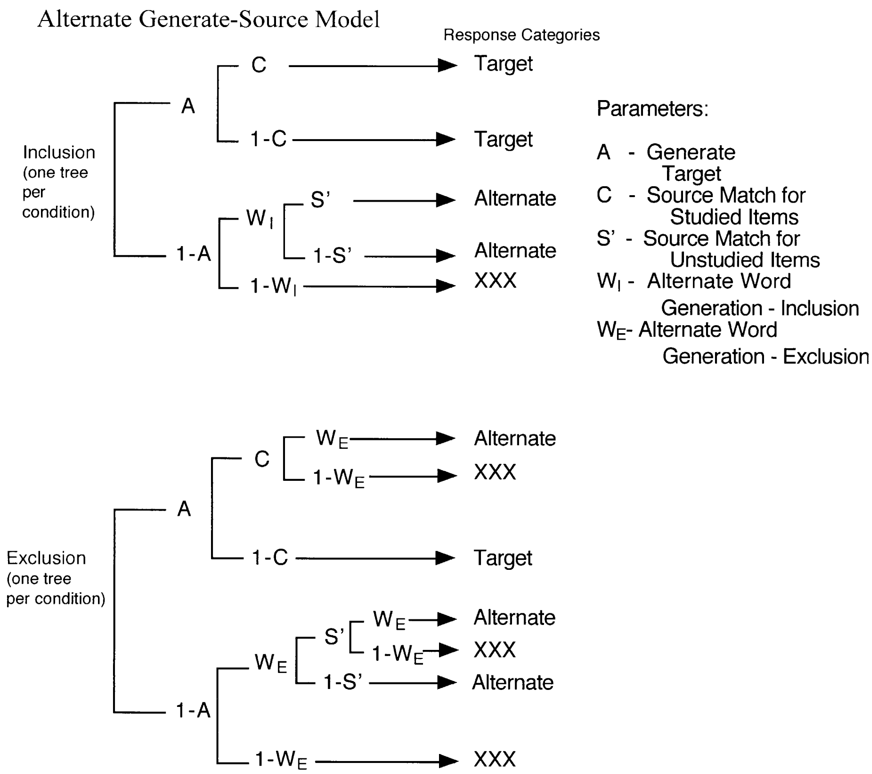

Fig. 3. Alternate form of the generate-source processing tree model from McBride and Dosher's (1999) and McBride et al.'s (2001) generate-recognize model of production tasks. *Note.* Unstudied items are also tested with the above trees. S′ is assumed for all scource matching for unstudied items.

pictures than words, which was predicted by TAP and is consistent with the results found by Wippich et al. (1998). The generate-source model also yields fairly stable estimates of conscious memory for pictures relative to words across tasks. For the Bodner et al. model, the ratios of C_P to C_W for the three experiments were 1.29, 1.28, and 1.51, respectively. These values are much more similar than those estimated by the direct-retrieval model and do not imply covariation of the conscious and automatic memory advantages for pictures. (The closer equivalence is also more consistent with the equivalent encoding conditions over experimental tasks. That is, the relative conscious advantage of picture to words follows from the relative strength of encoding, while the extent of expression may depend on the retrieval task.)

Since the direct-retrieval and generate-source models are nonnested, they are compared qualitatively. In the case of nonnested models, it is important to consider the validity of the patterns of the parameters being estimated. This is the primary method available for comparison of the models and the one being used here. Based on the patterns of results discussed above for the two models and the results of Bodner et al. (2000), the generate-source model appears favorable for these tasks and especially favorable for modeling processing on a stem completion task. A recent experiment (McBride, 2001) supports the validity of the generate-source model for stem completion results. After a study phase with picture and word items, participants performed stem completion trials, where they were to generate the first word that came to mind for the stem and then to determine the source for the item (picture, word, or new). Estimates of C and A for the generate-source model fit were consistent with estimates from Experiment 3; the A estimate for words was higher than that for pictures, but the C estimate for words was lower than that for pictures.

Despite the current results, it should be noted that the direct-retrieval model does provide good fits to data in some circumstances. For example, Jacoby (1998) found that with direct-retrieval instructions, the direct-retrieval model was fit well to stem completion data, while the generate-source model did not fit these data well. However, the use of direct-retrieval instructions does not appear to be sufficient to ensure good fits of the direct-retrieval model, as shown by the current results. One possible explanation for the superior fits of the generate-source over the direct-retrieval model in the current study could be the precise nature of the instructions that were used. Although direct-retrieval instructions were given in Experiments 2–4 of the current study, the exact wording of the instructions differed from those used by Jacoby (1998). The strength of the instructions may play a role in the type of strategy that participants employ in completing inclusion and exclusion tasks in the PDP. Given the inconsistent findings regarding the quality of fits for these models, the conditions in which participants adopt a direct-retrieval or generate-source strategy need to be investigated in future research.

10.2. Conceptual influences on automatic memory

Conceptual manipulations have previously been shown to influence implicit memory performance in some cases. Toth and Reingold (1996) presented a review of evidence of effects of conceptual factors on implicit memory tasks. Although such effects have been found in numerous studies, Toth and Reingold point out that very few of these effects have been confirmed using the PDP; therefore, the effects may be due explicit contamination on the implicit tests. For example, Mulligan (1997, 1998) has shown that an attention manipulation at study can affect implicit memory performance. However, Schmitter-Edgecombe (1999a) found no attention effects on *A* estimates in a PDP study. Further, Mulligan's (1997) research indicates that a conceptual implicit task is only affected by strong manipulations of attention, where explicit tasks are affected by much weaker manipulations. Parker, Gellatly, and Waterman (1999) also found effects of a context manipulation (a conceptual factor) on conceptual implicit tests (category production and general knowledge tasks), but not on perceptual implicit tests (fragment completion and anagram solution tasks). These results support a possible, but weak, link between conceptual processing and automatic memory for conceptual tasks.

Conceptual influence on automatic memory estimates may explain the higher *A* estimates for pictures than words from the current generate-source model fits to the category production task data in Experiment 4. Unlike the stem completion and picture identification tasks, the category production task has been shown to rely primarily on conceptual processing (Weldon & Coyote, 1996). If conceptual processing affects automatic forms of memory even slightly, these effects are more likely to be displayed on this task than on stem completion or picture identification.

10.3. Theories of picture superiority

In addition to support for the importance of the interaction of encoding and retrieval in accounting for picture superiority effects, these results confirm that pictures elicit more conceptual processing than words for all three tasks. Nelson (1979) described a model of picture and word encoding in which pictures directly activate a meaning code, while words activate a meaning code indirectly through a phonetic representation. This model of encoding is consistent with the greater conceptual processing for pictures claimed by Weldon and Roediger (1987). This model is also compatible with a TAP explanation of task performance. The sensory-semantic model of encoding suggested by Nelson (1979) distinguishes between conceptual and perceptual processing of pictures and words and is, therefore, compatible with the results found for the tasks in the current study. However, Nelson's theory does not make task-specific predictions in the way that TAP theory does. Transfer-appropriate processing theory assumes that an overlap in processing between study and test supports higher task performance and, therefore, allows for predictions based on the type of processing required by a task.

Dual-coding theory (Paivio, 1975, 1986, 1995), on the other hand, does not distinguish between conceptual and perceptual processing. The theory instead explains picture superiority as a product of the availability of multiple codes for pictures. Therefore, a dual-coding explanation does not predict task-specific access as does TAP, and dual-coding would have to be significantly elaborated to accommodate the results of these experiments. The current results, taken with previous evidence against dual-coding theory, indicate that this long-accepted theory may not be the best explanation of picture superiority. Since different tasks can result in different amounts of perceptual processing for pictures and words and can show a perceptual advantage for words or pictures depending on the stimulus given at test, a theory of picture superiority that can account for processing differences must be considered.

10.4. Conclusions

Overall, the current results are consistent with an encoding–retrieval interaction explanation of picture superiority. The logic of the explanation is that performance on a task depends both on the type and amount of processing engaged at both study and test. When the processing at study is reinstated at test, performance on implicit tests improves. Roediger and his colleagues (Roediger, 1990; Weldon & Roediger, 1987; Weldon et al., 1989) explain the picture superiority effect in this manner. They claim that pictures generally elicit more conceptual processing than words at study and that this is why they result in better performance than words on tasks that require conceptual processing (e.g., recall and recognition).

The results reported here for the three tasks generally support the explanation described by Weldon and Roediger (1987). If pictures elicit greater conceptual processing at study than words, they should have a conceptual advantage on a test requiring conceptual processing. In the current study, this is seen in higher C estimates for pictures than words on all tasks. According to the theory, perceptual processing, which influences estimates of automatic memory, should vary with the type of stimuli (words or pictures) and type of retrieval task. The perceptual estimates from the current data were as predicted for both stem completion and picture fragment identification tasks. For the direct-retrieval model, A estimates showed a perceptual advantage for pictures on the picture fragment task and an advantage for words on the word-stem completion task. The generate-source model showed the same results for A estimates. For the category production task, the generate-source model supported predictions for TAP theory: Pictures showed a conceptual advantage over words, as indicated by higher estimates (both C and A) for pictures than words.

Overall, the results support encoding–retrieval theories such as TAP, in which encoding and retrieval demands interact. Such theories provide a viable explanation of picture superiority effects. Future work should focus on developing methods to better specify the type and amount of processing required for particular tasks, knowledge important for making predictions based on TAP theory or other theories of this class.

Appendix A

A list of picture labels used for the stem completion task and the individual naming rates of each for the current study ($N = 11$) and Snodgrass and Vanderwart ($N = 42$) participants

Label	OURS (%)	S&V (%)	Label	OURS (%)	S&V (%)	Label	OURS (%)	S&V (%)
Anchor	100	93	Flower	100	93	Potato	100	90
Apple	100	98	Flute	82	88	Pumpkin	100	98

Appendix A (*continued*)

Label	OURS (%)	S&V (%)	Label	OURS (%)	S&V (%)	Label	OURS (%)	S&V (%)
Balloon	100	98	Football	100	100	Rabbit	82	100
Banana	100	100	Fork	100	100	Raccoon	91	79
Barn	70	69	Frog	100	100	Ring	100	98
Basket	100	90	Giraffe	100	95	Sandwich	100	100
Bear	100	88	Glasses	82	64	Scissors	100	98
Bell	100	100	Glove	82	98	Screw	70	93
Bird	100	88	Goat	91	86	Seal	100	88
Book	100	100	Gorilla	64	76	Shirt	82	100
Bottle	91	95	Grapes	100	90	Shoe	100	95
Bowl	91	95	Guitar	100	98	Skirt	82	98
Bread	100	83	Hammer	100	100	Sled	90	98
Broom	100	100	Hanger	91	86	Snake	100	98
Brush	100	83	Harp	73	93	Snowman	100	100
Butterfly	100	100	Heart	100	100	Sock	100	100
Camel	100	95	Helicopter	100	95	Spider	91	88
Candle	100	100	Horse	100	100	Spoon	100	98
Carrot	100	100	House	100	95	Squirrel	82	93
Caterpillar	91	79	Iron	100	95	Star	100	100
Celery	90	76	Kite	100	100	Strawberry	100	90
Chair	100	100	Knife	91	90	Sweater	82	83
Church	100	93	Lamp	100	93	Swing	73	95
Clown	100	95	Leaf	100	90	Table	100	95
Coat	100	79	Lock	90	88	Thimble	82	83
Comb	100	93	Monkey	100	95	Tiger	91	93
Corn	91	81	Moon	100	62	Tomato	100	88
Couch	73	67	Mushroom	100	98	Toothbrush	60	98
Crown	100	100	Necklace	100	60	Train	100	86
Desk	100	95	Needle	82	81	Tree	100	100
Donkey	82	86	Nose	100	98	Turtle	91	95
Door	100	98	Orange	70	81	Umbrella	100	100
Drum	100	98	Ostrich	70	86	Vest	100	98
Duck	100	95	Pants	91	88	Violin	73	86
Elephant	100	100	Peacock	73	79	Watch	100	90
Envelope	100	98	Pencil	100	100	Well	100	90
Fence	91	74	Pineapple	100	100	Wheel	100	93
Finger	91	71	Pipe	90	98	Whistle	100	100
Fish	100	100	Pitcher	73	88	Windmill	100	98
Flag	100	95	Plug	91	88	Wrench	82	76

Appendix B

A list of the stimuli used for the picture fragment identification task and correct naming rates of degraded items obtained during the pilot portion of the experiment ($N = 39$)

Item	% Naming	Item	% Naming	Item	% Naming	Item	% Naming
Barn	33.3	Flag	61.5	Needle	15.4	Spider	15.4
Basket	25.6	Flower	56.4	Nose	15.4	Strawberry	53.8
Bear	48.7	Flute	5.1	Orange	10.3	Sweater	43.6
Bread	25.6	Fly	46.2	Pants	38.5	Thimble	23.1
Car	48.7	Football	43.6	Peacock	5.1	Tomato	51.3
Caterpillar	12.8	Fork	30.8	Pencil	17.9	Toothbrush	12.8
Celery	25.6	Frog	28.2	Pineapple	33.3	Turtle	56.4
Church	25.6	Glove	61.5	Pitcher	48.7	Vest	38.5
Clown	7.7	Goat	43.6	Plug	5.1	Violin	41.0
Coat	38.5	Gorilla	35.9	Potato	5.1	Watch	10.3
Comb	10.3	Grapes	30.8	Raccoon	5.1	Well	35.9
Corn	10.3	Harp	33.3	Ring	20.5	Whistle	28.2
Couch	51.3	Helicopter	33.3	Sandwich	10.3	Wrench	17.9
Crown	55.3	House	59.0	Saw	30.8		
Desk	48.7	Iron	51.3	Screw	12.8		
Donkey	12.8	Knife	46.2	Seal	7.9		
Door	56.4	Ladder	12.8	Shirt	30.8		
Elephant	46.2	Lock	43.6	Skirt	20.5		
Fence	7.7	Monkey	7.7	Sled	17.9		
Finger	48.7	Necklace	17.9	Snake	28.2		

Appendix C

A list of picture labels and categories used for the category production task[7]

Four-legged animal	Tool	Vegetable
Camel (cat)	Ladder (level)	Asparagus (artichoke)
Giraffe (gopher)	Saw (screwdriver)	Carrot (corn)
Raccoon (rabbit)	Wrench (winch)	Potato (pepper)
Article in the kitchen	Article of clothing	Insect
Bowl (blender)	Glove (girdle)	Bee (beetle)
Pitcher (pot)	Pants (pajamas)	Caterpillar (cockroach)
Spoon (sifter)	Sweater (shoe)	Fly (flea)

[7] A possible alternate for each exemplar is provided in parentheses.

Appendix C (*continued*)

Furniture	Musical instrument	Vehicle
Couch (chair)	Bell (bassoon)	Bus (boat)
Desk (dresser)	Flute (fiddle)	Car (cart)
Table (television)	Harp (harpsichord)	Train (tank)
Part of the human body	Bird	Fruit
Ear (elbow)	Ostrich (owl)	Apple (apricot)
Finger (face)	Peacock (partridge)	Banana (berry)
Nose (neck)	Swan (swallow)	Pineapple (peach)
Office item	Bathroom item	
Envelope (eraser)	Comb (curlers)	
Pencil (pen)	Hairbrush (hairpin)	
Scissors (stapler)	Toothbrush (toilet)	

Appendix D

Details of the model fits for the alternative source model are given below by experiment. Parameter estimates are given for each item type with standard deviations in parentheses

Experiment 2—Picture fragment identification[a]

	C	A
Pictures	0.684 (0.032)	0.516 (0.020)
Words	0.527 (0.054)	0.318 (0.019)
Unstudied	0.001 (0.113)	0.206 (0.016)

[a] $G^2(3) = 0.92$; $W_I = 0.896$; $W_E = 0.886$; $S' = 0.799$.

Experiment 3—Stem completion[b]

	C	A
Pictures	0.863 (0.017)	0.531 (0.017)
Words	0.692 (0.023)	0.619 (0.016)
Unstudied	0.306 (0.050)	0.357 (0.016)

[b] $G^2(3) = 10.90$; $W_I = 0.735$; $W_E = 0.948$; $S' = 0.0001$.

Experiment 4—Category production[c]

	C	A
Pictures	0.762 (0.026)	0.572 (0.020)
Words	0.507 (0.040)	0.504 (0.020)
Unstudied	0.003 (0.080)	0.342 (0.019)

[c] $G^2(3) = 4.03$; $W_I = 0.769$; $W_E = 0.833$; $S' = 0.468$.

References

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

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57–86.

Bodner, G. E., Masson, M. E. J., & Caldwell, J. I. (2000). Evidence for a generate-recognize model of episodic influences on word-stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 267–293.

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 404–416.

Buchner, A., & Erdfelder, E. (1996). On assumptions of, relations between, and evaluations of some process dissociation measurement models. *Consciousness and Cognition, 5*, 581–594.

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 124*, 137–160.

Challis, B. H., & Brodbeck, D. R. (1992). Level of processing affects priming in word fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 595–607.

Cowan, N., & Stadler, M. A. (1996). Estimating unconscious processes: Implications of a general class of models. *Journal of Experimental Psychology: General, 125*, 195–200.

Craik, F. I. M., Moscovitch, M., & McDowd, J. M. (1994). Contributions of surface and conceptual information to performance on implicit and explicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 864–875.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 531–547.

Hamann, S. B. (1990). Level-of-processing effects in conceptually driven implicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 970–977.

Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods, Instruments, and Computers, 31*, 220–234.

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 3–26.

Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language, 35*, 32–52.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language, 29*, 433–454.

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

Jones, G. V. (1987). Independence and exclusivity among psychological processes: Implications for the structure of recall. *Psychological Review, 94*, 229–235.

Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General, 122*, 462–467.

McBride, D. M. (2001). The effect of retrieval orientation on fits of direct-retrieval and generate-source models (submitted).

McBride, D. M., & Dosher, B. A. (1999). Forgetting rates are comparable in conscious and automatic memory: A process-dissociation study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 583–607.

McBride, D. M., Dosher, B. A., & Gage, N. M. (2001). A comparison of forgetting for conscious and automatic memory processes in word fragment completion tasks. *Journal of Memory and Language, 45*, 585–615.

Mecklenbraeuker, S., Wippich, W., & Mohrhusen, S. H. (1996). Conscious and unconscious influences of memory in a conceptual task: Limitations of a process-dissociation procedure. *Swiss Journal of Psychology, 55*, 34–48.

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

Mulligan, N. W. (1997). Attention and implicit memory tests: The effects of varying attentional load on conceptual priming. *Memory & Cognition, 25*, 11–17.

Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 27–47.

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. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 485–497.

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

Paivio, A. (1975). Coding distinctions and repetition effects in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 179–214). New York: Academic Press.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.

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

Paivio, A. (1995). Imagery and memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 977–986). Cambridge, MA: MIT Press.

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

Parker, A., Gellatly, A., & Waterman, M. (1999). The effect of environmental context manipulation on memory: Dissociation between perceptual and conceptual implicit tests. *European Journal of Cognitive Psychology, 11*, 555–570.

Rajaram, S., & Roediger, H. L., III (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 765–776.

Reingold, E. M., & Wainwright, M. J. (1996). Response bias correction in the process dissociation procedure: A reevaluation? *Consciousness and Cognition, 5*, 595–603.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*, 475–543.

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

Roediger, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.

Roediger, H. L., III, Weldon, M. S., Stadler, M. A., & Riegler, G. H. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1251–1269.

Rothkegel, R. (1999). AppleTree: A multinomial processing tree modeling program for Macintosh computers. *Behavior Research Methods, Instruments, and Computers, 31*, 696–700.

Russo, R., Cullis, A. M., & Parkin, A. J. (1998). Consequences of violating the assumption of independence in the process dissociation procedure: A word fragment completion study. *Memory & Cognition, 26*, 617–632.

Schacter, D. L. (1987). Implicit memory: History and status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501–518.

Schmitter-Edgecombe, M. (1999a). Effects of divided attention on perceptual and conceptual memory tests: An analysis using a process dissociation approach. *Memory & Cognition, 27*, 512–525.

Schmitter-Edgecombe, M. (1999b). Effects of divided attention and time course on automatic and controlled components of memory in older adults. *Psychology and Aging, 14*, 331–345.

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

Srinivas, K., & Roediger, H. L., III (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language, 29*, 389–412.

Toth, J. P., & Reingold, E. M. (1996). Beyond perception: Conceptual contributions to unconscious influences of memory. In G. Underwood (Ed.), *Implicit cognition* (pp. 41–84). New York: Oxford University Press.

Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 290–303.

Wainwright, M. J., & Reingold, E. M. (1996). Response bias correction in the process dissociation procedure: Approaches, assumptions, and evaluation. *Consciousness and Cognition, 5*, 232–254.

Weldon, M. S. (1993). The time course of perceptual and conceptual contributions to word fragment completion priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1010–1023.

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, and Cognition, 22*, 670–686.

Weldon, M. S., & Jackson-Barrett, J. L. (1993). Why do pictures produce priming on the word fragment completion test? A study of encoding and retrieval factors. *Memory & Cognition, 21*, 519–528.

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.

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*, 33–46.