

Using What We Know: Consequences of Intentionally Retrieving Gist Versus Item-Specific Information

Wilma Koutstaal

University of Minnesota and University of Reading

Margaret Cavendish

University of Reading

The effect of a prior gist-based versus item-specific retrieval orientation on recognition of objects and words was examined. Prior item-specific retrieval increased item-specific recognition of episodically related but not previously tested objects relative to both conceptual- and perceptual-gist retrieval. An item-specific retrieval advantage also was found when the stimuli were words (synonyms) rather than objects but not when participants overtly named objects during gist-based recognition testing, which suggests that they did not always label objects under general gist-retrieval instructions. Unlike verbal overshadowing, labeling objects during recognition attenuated (but did not eliminate) test- and interference-related forgetting. A full understanding of how retrieval affects subsequent memory, even for events or facts that are not themselves retrieved, must take into account the specificity with which that retrieval occurs.

Keywords: episodic memory, categorical memory, retrieval-induced forgetting, problem solving, thinking

A central and well-documented observation about human memory and knowledge is that humans can represent and access information at varying levels of specificity or grain. People may choose (or find themselves operating with) representations at any of a number of levels of abstraction, with differing degrees of tight, strict, or verbatim correspondence to the original perceptual or conceptual content of their experiences. This is true both for individual spatiotemporal events (episodic and autobiographical memory; e.g., S. J. Anderson & Conway, 1993; Brainerd & Reyna, 1998; Conway, 1992; Goldsmith, Koriat, & Weinberg-Eliezer, 2002) and for general facts and knowledge (semantic memory; e.g., Murphy & Smith, 1982; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Tanaka & Taylor, 1991; see Cohen, 2000, and Murphy, 2002, for reviews). The end points of the specificity continuum have been described with various (not necessarily interchangeable) terms, with the more specific end denoted as *item specific*, *exemplar based*, or *verbatim* and the more abstract end

denoted as *category based*, *gist based*, or *meaning based* (e.g., Brainerd & Reyna, 1990; Koutstaal & Schacter, 1997; Murphy, 2002). A person's cognitive-perceptual position on this specificity continuum for any given task or situation often is a key factor in the types of decisions, judgments, and errors he or she may make (Ewa, 2002; Koutstaal, 2003; Koutstaal, Schacter, Galluccio, & Stofer, 1999; Pansky & Koriat, 2004; Reyna, 2004; Yaniv & Foster, 1995) and has been found to be influenced by such diverse factors as age (Brainerd & Gordon, 1994; Koutstaal & Schacter, 1997; Micco & Masson, 1992; Tun, Wingfield, Rosen, & Blanchard, 1998), time since encoding (Brainerd & Reyna, 1990; Dorfman & Mandler, 1994), an orientation toward the avoidance of particular types of errors (Koriat & Goldsmith, 1996; Yaniv & Foster, 1995), level of expertise or training (Johnson & Mervis, 1997; Kulatunga-Moruzi, Brooks, & Norman, 2001; Tanaka & Taylor, 1991), and affective state (e.g., Raes et al., 2005; Ramponi, Barnard, & Nimmo-Smith, 2004; Williams & Broadbent, 1986).

A second, equally central and well-documented observation about human knowledge and memory is that the act of retrieving and using information, through intentional or incidental recall or reminding, significantly affects subsequent memory. Depending on circumstances, an act of retrieval may substantially enhance the later ease of retrieval of a given event (e.g., Allen, Mahler, & Estes, 1969; Hogan & Kintsch, 1971; Koutstaal, Schacter, Johnson, Angell, & Gross, 1998; Linton, 1975, 1978), impede the retrieval of related events or knowledge (e.g., M. C. Anderson, Bjork, & Bjork, 1994; Bäuml, 2002; Koutstaal, Schacter, Johnson, & Galluccio, 1999; Perfect et al., 2004), or distort and bias subsequent retrieval (e.g., Roediger, Jacoby, & McDermott, 1996; Schooler, Foster, & Loftus, 1988; Schreiber & Sergeant, 1998; Tuckey & Brewer, 2003).

Notably, however, despite the extensive and wide-ranging investigations of both of these simple observations about human

Wilma Koutstaal, Department of Psychology, University of Minnesota, and Department of Psychology, University of Reading, Reading, England; Margaret Cavendish, Department of Psychology, University of Reading.

This research was supported by Economic and Social Research Council Grants R000223749 and R000220429 and by the Nuffield Undergraduate Summer Research Program.

We thank Wendy Gibbons (Master of Science Student Placement Program), Eva Moran (supported by the Economic and Social Research Council), Nicola Perkins (Nuffield Undergraduate Summer Research Program), Elena Sakkalou (Master of Science Student Placement Program), and Amanda Townson for their help with the preparation of materials and testing of participants for Experiments 1 and 3. We also thank Judi Ellis for her invariably thoughtful advice.

Correspondence concerning this article should be addressed to Wilma Koutstaal, Department of Psychology, University of Minnesota, 247 Elliott Hall, 75 East River Road, Minneapolis, MN 55455. E-mail: kouts003@umn.edu

cognition, the interconnection of these two observations has scarcely been examined. That is, how does the level of specificity of retrieval at one point in time affect subsequent retrieval at either the same or a different level of retrieval? Can people adaptively switch between levels of retrieval as required by changing circumstances, or does an earlier adopted level of specificity “carry over” (with either beneficial or harmful effects to judgment or performance) to their later attempts to use information at a different level of specificity?

Memory research has primarily investigated access to specific representations and has treated gist-based processing as a source of error (Goldsmith et al., 2002; Koutstaal, 2003). Yet judgment, reasoning, and problem solving often require a meaning-based, category-based, or gistlike focus. Judgments and tasks that investigators have argued draw on gist-based representations include categorical inference, analogical transfer, transitive inference, class inclusion, metaphorical interpretation, and pragmatic inference. In these and other tasks, researchers have argued that forgetting of details may be adaptive because it increases abstractness and flexibility; reliance on gist also may enhance performance by circumventing sources of memorial and processing failures (e.g., Brainerd & Reyna, 1990, 1992; Catrambone & Holyoak, 1989; Gick & Holyoak, 1983; Matthews, 1991).

Many studies have demonstrated that there may be detrimental effects of prior retrieval of some items on the accessibility of other related items, both for recently learned episodic information (e.g., Ciranni & Shimamura, 1999; Koutstaal, Schacter, Johnson, & Galluccio, 1999; Mueller & Watkins, 1977; Perfect et al., 2004; Shaw, Bjork, & Handal, 1995) and for long-term semantic or propositional knowledge (e.g., M. C. Anderson & Bell, 2001; M. C. Anderson et al., 1994; J. Brown, 1968; Karchmer & Winograd, 1971). However, the effects of such prior retrieval on access to knowledge at different levels of specificity have not been assessed or evaluated. Nevertheless, it is clear that people can (and often must) recall information at different levels of specificity. Further, although situational and individual factors may promote reliance on either gist or item-specific information, recent studies using a variety of procedures and neuropsychological populations have suggested that, in many instances, both forms of information are at least potentially available for retrieval and may interact with or play off one another (e.g., Koutstaal, Verfaellie, & Schacter, 2001; Reyna & Brainerd, 1995).

The primary purpose of the experiments reported in this article is to examine the impact of an explicitly (intentionally) adopted gist-based versus item-specific retrieval orientation on later knowledge availability. How does a gist-based retrieval focus—often needed for judgment and reasoning—affect later accessibility of knowledge? What are the consequences of earlier retrieval probes of an episode made at one level of specificity for the success of subsequent attempts at a different level of specificity? Each of the experiments used an episodic recognition paradigm wherein recognition at either an item-specific or a gist- or category-based level was probed successively for a large number of stimuli encoded in a single “episode.” In all cases, the effects of prior retrieval specificity were assessed on the ability to recognize other stimuli from the initial episode that had not been previously tested.

Experiment 1: Conceptual Focus

To evaluate the effects of adopting a retrieval orientation toward more categorical versus item-specific aspects of an episode, when particular events within that episode were not themselves directly queried, we administered two initial test blocks probing either conceptual or item-specific recognition to induce a given retrieval set. These blocks then were followed by a further critical test block. Participants first encountered a large set of items (incidental encoding) and then were given three separate episodic recognition test blocks, each composed of nonoverlapping items (i.e., no items were tested repeatedly). The key point of interest is performance on the third test; the first two tests acted to induce a particular retrieval set. Retrieval set was manipulated between subjects. There were four conditions: consistent item-specific or identical recognition (identical, identical, identical [III]), consistent conceptual recognition (conceptual, conceptual, conceptual [CCC]), and two forms of changed set involving a change from identical to conceptual recognition (identical, identical, conceptual [IIC]) or from conceptual to identical recognition (conceptual, conceptual, identical [CCI]). Does the induction of a consistent retrieval orientation toward a more generic or gist-based level render it more difficult to subsequently access item-specific details about other (not previously retrieved) items?

The identical recognition test involved simple old–new episodic recognition. However, in addition, the instructions to participants strongly emphasized that they were to call an item “old” only if it was identical to an item they had encountered previously (see Figure 1). By contrast, the conceptual or meaning-based recognition test also involved episodic recognition instructions, but the type of judgment requested was different. Participants were asked to designate an item as old either if it was identical to one presented earlier or if it was similar in meaning to or conceptually related to an item presented earlier (see Figure 1; cf. Brainerd & Reyna, 1998; Reder, Wible, & Martin, 1986).

Method

Participants. The participants were 72 undergraduates from the University of Reading, Reading, England, recruited through posted flyers. Their average age was 19.71 years ($SD = 1.67$); they had, on average, 14.74 years ($SD = 1.65$) of formal education. All participants reported normal or corrected-to-normal vision and color vision and were native speakers of English. Additionally, all participants were screened for depression with the Brief Symptom Inventory (Derogatis & Melisaratos, 1983), and those who had scores of 11 or higher on the depression items were replaced. Individuals participated for course credit or for a small monetary remuneration. A block-random procedure was used to assign an equal number of participants to each of the four between-subjects conditions: CCC (all conceptual recognition tests), III (all identical recognition tests), CCI (conceptual recognition followed by identical recognition), or IIC (identical recognition followed by conceptual recognition).

Stimulus materials. The stimuli were colored photographs or detailed line drawings of common objects or animals (see Figure 1). There were a total of 240 object pairs (plus practice items); 120 items (1 each from 120 pairs) were presented at study. Each of the three memory tests consisted of 120 items, including the presented and nonpresented exemplars of one set of 40 items (i.e., 40 same and 40 different exemplars), plus 40 new exemplars, presented in a pseudorandomly intermixed order. The items presented at study versus test and the type of test on which they appeared were counterbalanced across participants, and any given item was tested

Study Phase	Test Phase			
	Stimulus	Recognition Type	Stimulus	Correct Response
	Item-Specific		New	Unrelated
	Item-Specific		New	Different Exemplar
	Item-Specific		Old	Same Exemplar
	Category-based		Old	Different Exemplar
	Category-based		New	Unrelated
	Category-based		Old	Same Exemplar

Figure 1. Examples of the stimuli and recognition tasks. Although they are shown here in gray scale, stimuli in the experiment were shown in color. See the text for details.

only once per participant. Stimuli were presented on a Macintosh computer with a color monitor, and participants entered their recognition responses using the keyboard.

Procedure. All participants were tested individually. After they provided informed consent, the experiment consisted of four separate phases: an encoding phase followed by three episodic recognition tests. In the encoding phase, participants performed a size-judgment task in which they indicated whether the real-world referent of the object shown was larger than a 13-in. (33-cm) box (an example box was provided). This incidental encoding task was followed by three successive episodic recognition tests, with the particular instructions for each recognition test determined by the participant's condition assignment (i.e., CCC, III, CCI, or IIC) and the two forms of instructions, as described earlier. Examples of identical and conceptually related items were provided. After completing all of the episodic tests, participants were debriefed.

Results

The analyses for this and each of the following experiments concentrate on two primary comparisons: (a) the effects of a sustained prior retrieval orientation toward item-specific versus gist-based information on the accuracy of item-specific recognition, which we evaluated by contrasting *item-specific recognition* on the third test in the CCI versus III conditions, and (b) the effects of a sustained prior retrieval orientation toward item-specific versus gist-based information on the accuracy of gist-based recogni-

tion, which we evaluated by contrasting *conceptual recognition* on the third test in the IIC versus CCC conditions. In each case, we computed analyses for three measures of sensitivity (A') and corresponding measures of response bias ($B_{D'}$), computed as described in Koutstaal and Schacter (1997). A' can vary between 0.00 and 1.00, with higher values indicating greater sensitivity and chance performance being .50. $B_{D'}$ can vary between -1.00 and 1.00 , with negative values representing more liberal and positive values more stringent responding. The three measures of sensitivity and response bias were based on (a) a comparison of hits for same exemplars versus false alarms to novel (unrelated) items, termed A'_{-novel} and $B_{D'_{-novel}}$, respectively; (b) a comparison of hits for same exemplars versus false alarms to related or different exemplars, termed $A'_{-related}$ and $B_{D'_{-related}}$, respectively; and (c) a comparison of "hits" for different exemplars versus false alarms to novel (unrelated) items, termed A'_{-gist} and $B_{D'_{-gist}}$, respectively. Both A'_{-novel} and $A'_{-related}$ provide measures of item-specific memory. However, A'_{-novel} also may reflect the use of gist- or category-based information (e.g., responding "yes" on the basis that one generally remembers having encountered an object of a given kind, such as an umbrella, without remembering any further details). By contrast, $A'_{-related}$ is a more stringent measure of item-specific memory because it requires fine-grained discrimination between studied items and distractor items from the same

object category. $A'_{\text{-gist}}$ reflects responding to object category information. Figure 2 (upper panel) presents these three sensitivity measures for the final test for the four conditions (III, CCI, CCC, and CCI); Table 1 presents the average probability of “old” responses on the final test separately for same exemplars, different exemplars, and new (unrelated) items.¹

Effects of prior retrieval orientation on item-specific recognition. Individuals who consistently engaged in item-specific retrieval across the three tests showed higher item-specific sensitivity on the final test ($A'_{\text{-novel}} = .90$) than did individuals who had first engaged in conceptual recognition for the first two tests and then switched to item-specific retrieval on the third test ($A'_{\text{-novel}} = .86$). A one-way between-subjects analysis of variance contrasting $A'_{\text{-novel}}$ on the third test for the III versus the CCI groups showed a significant detrimental effect of the prior conceptual retrieval orientation, $F(1, 34) = 4.18$, $MSE = .002$, $p = .049$, $d = 0.70$. There was a similar, slightly more marked detrimental effect of prior conceptual ($A'_{\text{-related}} = .80$) compared with prior item-specific ($A'_{\text{-related}} = .86$) retrieval orientation when we compared hits to same exemplars versus the related lures, $F(1, 34) = 5.42$, $MSE = .007$, $p = .03$, $d = 0.80$. The two conditions did not differ in gist sensitivity ($F < 1$) or on any of the measures of response bias ($F_s < 2$).

Additional analyses showed that it was unlikely that these differences in sensitivity reflected initial differences in the groups, because the item-specific recognition performance of the III and IIC groups—who were assigned to conditions according to exactly the same procedure as for the III and CCI groups—did not differ on the very first test (i.e., before the retrieval manipulation; $F_s < 1.2$). Similar findings of no significant initial *premanipulation* differences at Test 1 were obtained for all sensitivity and response bias comparisons in this and the following experiments.

Effects of prior retrieval orientation on gist-based recognition. When we contrasted the conceptual recognition performance of individuals who had engaged in conceptual recognition throughout the testing session versus those who had first performed item-specific recognition and then switched to conceptual recognition on the final test, we also found a conceptual recognition advantage for the IIC group ($A'_{\text{-novel}} = .90$) relative to the CCC group ($A'_{\text{-novel}} = .86$), $F(1, 34) = 4.85$, $MSE = .004$, $p = .03$, $d = 0.76$. We observed a similar pattern for the measure of $A'_{\text{-gist}}$, which reflects greater sensitivity to the different exemplars on the conceptual task following identical ($A'_{\text{-gist}} = .88$) rather than conceptual ($A'_{\text{-gist}} = .84$) retrieval; however, this pattern was not shown as consistently across all participants, $F(1, 34) = 2.62$, $MSE = .006$, $p = .11$, $d = 0.56$, observed power = .50. There were no effects on $A'_{\text{-related}}$ or any of the bias measures ($F_s < 1$).

Discussion

This experiment showed that sustained efforts to retrieve encoded information from a categorical or meaning-based focus impeded participants' ability to recognize other episodically related items at a specific level (CCI vs. III comparison; significant adverse effect observed on both the $A'_{\text{-novel}}$ and $A'_{\text{-related}}$ measures of sensitivity). Sustained categorical retrieval also impeded subsequent categorical level recognition compared with first adopting an item-specific focus and then attempting category-based recognition (CCC vs. IIC comparison; significant adverse effect for same exemplars, as shown in the $A'_{\text{-novel}}$ analyses).

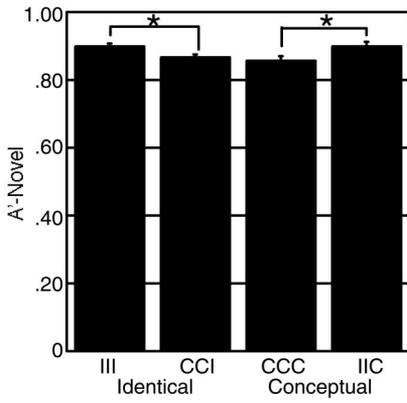
These outcomes are a powerful demonstration of the detrimental consequences of a category-based episodic retrieval focus on memory for other events from the same episodic context, because several aspects of the experiment should have operated against the emergence of such an effect. For example, the experiment involved recognition testing (many contextual effects are more pronounced in recall than in recognition, which suggests that the effect is very robust), and the stimuli were pictures of common objects that were rich in perceptual detail that could have reinstated the earlier encounter. These findings point to the beneficial effects of an item-specific focus during knowledge access (cf. Medin & Ross, 1989) and are in tension with observations and accounts pointing to a believed facilitative role of a meaning-based abstract or conceptual focus in many cognitive tasks, such as problem solving, reasoning, and comprehension (e.g., Brainerd & Reyna, 1990). These results suggest that an abstract conceptual focus may have unintended detrimental effects—leading to decreased availability of episodically related (but nonretrieved) item-specific and conceptual knowledge.

The observed harmful consequences of a categorical retrieval focus raise a number of important further questions. What processes or mechanisms are involved in the development of the detrimental effects arising from a sustained meaning- or gist-based retrieval orientation? In addition (closely related to the question of the underlying processes and mechanisms), how general is the effect? For example, are there similar detrimental effects on accessibility of knowledge following a sustained gist-based retrieval orientation when the stimuli are themselves more abstract and less perceptual in nature (words rather than pictures)? Would one expect a similar outcome if the similarity-based abstraction encouraged during gist-based retrieval involved visual-perceptual similarity rather than object-level categorical information—for example, if the gist-based orientation was founded on similarity of perceptual form rather than conceptual or semantic similarity?

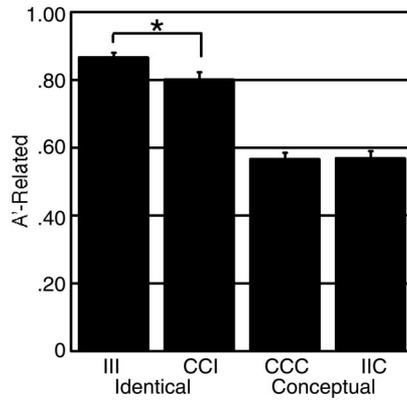
The theoretical approach to these findings is still tentative, but we test and develop it through the experiments described in the following sections. Two sets of prior findings are especially relevant. The first were reported by Whittlesea, Brooks, and Westcott (1994) in an article tellingly titled “After the Learning Is Over: Factors Controlling the Selective Application of General and Particular Knowledge.” Across a series of five experiments, each using a domain of artificial words, these investigators demonstrated that when and whether newly acquired item- versus class-level knowledge was used depended on more than the direct demands of the task at hand. A number of other factors, such as whether, within the tasks, class-level knowledge was made rela-

¹ On the conceptual tests, $A'_{\text{-related}}$ values greater than .50 indicate higher category- or gist-based recognition for same than for different exemplars, indicating an (unintentional) perceptual reinstatement advantage on the meaning-based test. Analyses (one-sample t tests against an expected population value of .50) showed significantly greater sensitivity to same than to different exemplars in all four experiments reported here, with a significant same-exemplar advantage observed in 17 out of the 18 conceptual tests using object stimuli and all 6 of the 6 conceptual tests using word stimuli. This consistent same- over different-exemplar conceptual recognition advantage replicates that found previously (Koutstaal, 2003) and may partially reflect what Brainerd and Reyna (1993) termed a “verbatim exit bias” (p. 52), or a tendency of individuals to make similarity judgments when they should be processing meaning (rather than surface form).

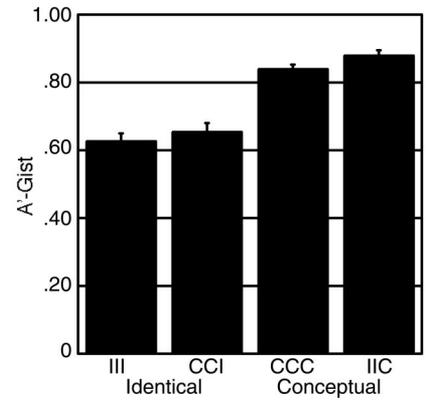
A'-Novel
Experiment 1



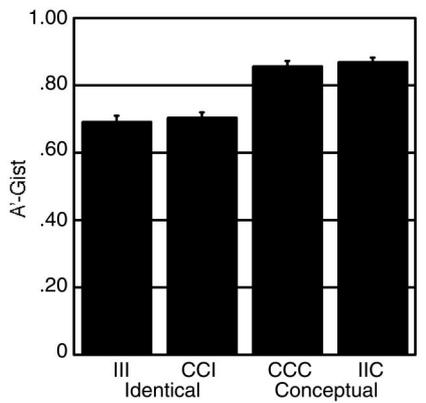
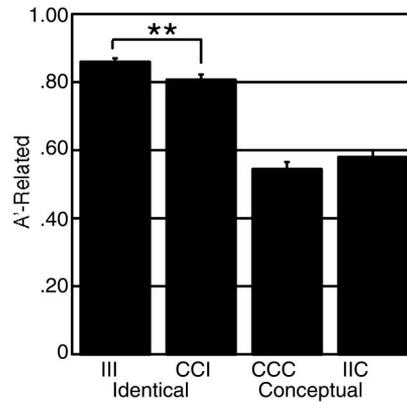
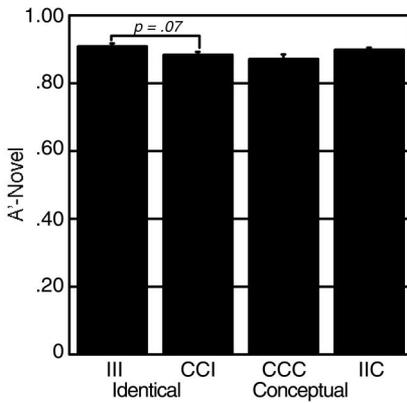
A'-Related



A'-Gist



Experiment 2



Experiment 4

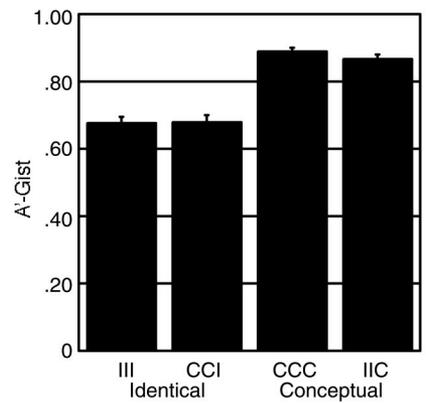
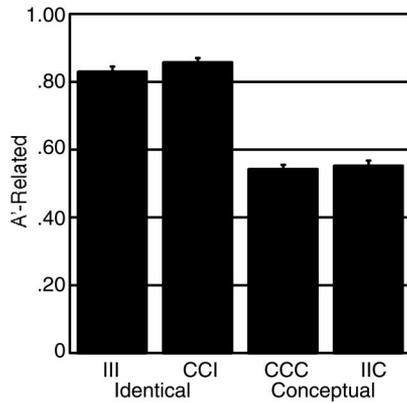
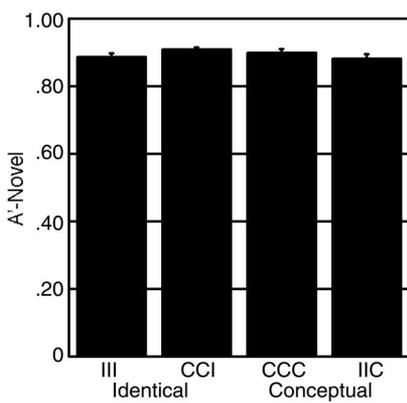


Table 1
Mean Proportion of “Old” Responses for the Final Test, Shown Separately for Groups Given Final Item-Specific (III, CCI) and Final Conceptual (CCC, IIC) Recognition Tests

Group	Item-specific recognition			Group	Category-based recognition		
	Same	Different	New		Same	Different	New
Experiment 1: Conceptual focus							
III	.73	.14	.07	CCC	.75	.71	.20
CCI	.64	.19	.09	IIC	.85	.80	.17
Experiment 2: Perceptual focus							
III	.74	.17	.04	CCC	.82	.78	.21
CCI	.65	.20	.05	IIC	.83	.76	.16
Experiment 3: Words rather than pictures							
III	.69	.31	.34	CCC	.64	.49	.38
CCI	.64	.36	.31	IIC	.65	.46	.39
Experiment 4: Conceptual focus with verbal labeling							
III	.68	.17	.06	CCC	.83	.80	.15
CCI	.73	.15	.04	IIC	.81	.76	.17

Note. Groups differed in the nature of the three successive recognition tests that they received: III = identical, identical, identical; IIC = identical, identical, conceptual; CCC = conceptual, conceptual, conceptual; CCI = conceptual, conceptual, identical. Same = same exemplar; Different = different exemplar; New = unrelated items. Under item-specific recognition, “old” responses to different exemplars are false recognition responses; under category-based recognition “old” responses to different exemplars are correct (gist-based) responses.

tively more accessible (through a “class label first” manipulation) or, instead, item knowledge was accessed more quickly (through an “item first” manipulation) influenced whether general versus particular knowledge was brought to bear.

The outcomes of the current experiment also have a family resemblance to work on *verbal overshadowing* (Schooler & Engstler-Schooler, 1990). In verbal overshadowing, recognition of visual or other complex perceptual stimuli (e.g., unfamiliar faces) may be impeded by the attempt, at the time of encoding, to provide a verbal description of the stimuli. Detrimental effects of verbal descriptions have been observed not only for stimuli that were themselves described but also for items (e.g., other faces) that were presented in the same episodic context as was the described stimulus (C. Brown & Lloyd-Jones, 2002; Dodson, Johnson, & Schooler, 1997; Westerman & Larsen, 1997). The generalized detrimental effects of verbalization have led to the proposal that recognition performance may be adversely affected by a more

general “retrieval-inappropriate” shift in processing orientation—that is, *transfer-inappropriate retrieval* (Fallshore & Schooler, 1995; Schooler, 2002; also cf. Macrae & Lewis, 2002). These findings from the verbal overshadowing literature mesh with those of the current experiment in suggesting that recent retrieval and processing provide an important and sometimes detrimental context for subsequent retrieval not only for items that are themselves retrieved but for other (episodically related) items as well; qualitative shifts in the types of information accessed at one point in time influence what then can be readily accessed. (Note that although, unlike in verbal overshadowing, no explicit verbal labeling was required during conceptual recognition, covert labeling might have occurred. We examine the effects of overt labeling in Experiment 4.)

Experiment 2: Perceptual Focus

Were the detrimental effects of a categorical focus found in Experiment 1 mediated by an increased emphasis on language or verbal labels? The requirement to make meaning-based compared with item-specific recognition judgments in Experiment 1 might have required (or encouraged) participants to recode the stimuli into lexical or phonological form. Such recoding might facilitate access to earlier encountered material from a categorical (meaning-based) perspective. For example, on being shown an umbrella, participants might name the object and ask themselves, “Did I see an instance of an umbrella?” rather than trying to recollect a specific pictorial representation of a previously encountered umbrella. By contrast, the item-specific retrieval focus might have encouraged processes involving object representations more closely tied to the original pictorial information. For the item-specific recognition probe, there might have been no need to attempt to recast the stimulus into a more abstract, possibly lexical form, because participants could best meet the task demands by considering the exact stimulus currently presented.

The aim of Experiment 2 is to test the effects of a sustained category-based retrieval orientation on later knowledge accessibility when the orientation requires a visual-perceptual rather than a semantically or conceptually based categorical focus. Rather than being asked whether they had encountered a conceptually similar object, participants in the visual-perceptual condition were asked to decide whether they earlier had been shown an object that was similar in shape to the current object. The instructions emphasized that participants should rely on visual-perceptual images rather than conceptual comparisons or labels.

If a general categorical orientation drives the detrimental meaning-focus effect (regardless of whether the categories are defined perceptually or semantically), then those harmful effects also should be observed for the perceptually based instructional set. However, if the detrimental gist-focus effect derives from a shift

Figure 2 (opposite). Measures of sensitivity ($A'_{\text{-novel}}$, $A'_{\text{-related}}$, and $A'_{\text{-gist}}$) for Experiments 1, 2, and 4, all of which used object stimuli. Sensitivity measures are shown separately for each of the four conditions, in which the type of recognition probe either remained constant (III = identical, identical, identical; CCC = conceptual, conceptual, conceptual) or changed on the final test (IIC = identical, identical, conceptual; CCI = conceptual, conceptual, identical). The measures of sensitivity shown are for the final (third) test in each condition; the prior two tests established an orientation to either item-specific or conceptual information. A' of .50 represents chance performance. Error bars show the standard error of the mean. * $p < .05$. ** $p < .01$.

toward verbal information processing (recoding the remembered stimuli into verbal format, somewhat akin to verbal overshadowing), then the perceptually focused classification instructions should ameliorate or even eliminate the detrimental category-focus effect.

Method

Participants. Participants were recruited and screened in the same manner as for Experiment 1, except that (a) a total of 96 participants were tested and (b) 3 individuals, although fluent in English, were not native speakers. The mean age of the participants was 19.20 years ($SD = 1.42$); they had, on average, 13.79 ($SD = 0.92$) years of formal education.

Stimuli. The stimuli were identical to those used in Experiment 1.

Procedure. Except for the change in the instructions for the meaning-based episodic recognition tests, the procedure was identical to that used in Experiment 1. In addition, 24 participants were tested in each of the four conditions. For the perceptually based recognition task, participants were told that they would be shown items that were exactly the same as the ones they were shown earlier and items that were perceptually similar to previously shown items (several examples were provided). Participants were asked to respond “old” to both “the same item that was shown before and the perceptually similar item” and were explicitly instructed to “make your judgment based on the shape of the object, rather than giving the shape a name.”

Results

Effects of prior retrieval orientation on item-specific recognition. Despite the change in instructions, there was again a trend toward an item-specific retrieval advantage, with sustained item-specific retrieval across the three tests leading to higher item-specific retrieval on the final test ($A'_{\text{-novel}} = .91$) than initial (perceptually focused) categorical retrieval followed by a switch to item-specific retrieval ($A'_{\text{-novel}} = .88$), $F(1, 46) = 3.53$, $MSE = .002$, $p = .07$, $d = 0.55$, observed power = .59. This prior item-specific ($A'_{\text{-related}} = .86$) compared with categorical ($A'_{\text{-related}} = .81$) retrieval advantage was again reliable and more strongly apparent on the $A'_{\text{-related}}$ measure, $F(1, 46) = 7.47$, $p = .009$, $d = 0.81$ (see Figure 2, middle panel). There were no effects on the gist measure of sensitivity ($F < 1$) or any measures of bias ($F_s < 2$).

Effects of prior retrieval orientation on gist-based recognition. In contrast to the preserved item-specific retrieval advantage on item-specific recognition, the item-specific retrieval advantage for categorical recognition appeared to be less uniformly shown under the perceptually focused instructions. For each sensitivity measure, IIC still showed a numerical advantage over CCC, but these differences were not significant (for same exemplars, $A'_{\text{-novel}} = .90$ vs. $.87$, $F = 2$; $A'_{\text{-related}} = .58$ vs. $.55$, $F = 1.5$; $A'_{\text{-gist}} = .87$ vs. $.86$, $F < 1$). (Note, however, that these patterns, although less consistent, were in the same direction as those shown in Experiment 1. Additionally, as we report in the following section, analyses on the combined results of Experiments 1 and 2 still showed a significant advantage of IIC over CCC for the $A'_{\text{-novel}}$ measure, with no Condition \times Experiment interaction.) There were no effects on any bias measures ($F_s < 1$).

Discussion

In this experiment, we manipulated participants' gist-based retrieval orientation to emphasize visual-perceptual form information (object shapes) rather than lexical or conceptual information. Despite these altered instructions, the comparison of identical recognition on the key third test in the III versus CCI conditions

revealed that there was again, as in Experiment 1, a substantial recognition advantage for the item-specific orientation. Additional analyses that combined the results from the two experiments showed that there was a robust and significant advantage of III over CCI for both the $A'_{\text{-novel}}$ sensitivity measure, $F(1, 80) = 7.71$, $MSE = .002$, $p = .007$, and the $A'_{\text{-related}}$ sensitivity measure, $F(1, 80) = 12.85$, $MSE = .005$, $p = .0006$, with no Condition \times Experiment interactions ($F_s < 1$).

In contrast to the identical recognition results, the specificity advantage on the final conceptual recognition test in the comparison of IIC versus CCC was no longer reliably shown. Nonetheless, analyses combining the two experiments showed that there was also a significant specificity advantage on the final conceptual recognition test in the comparison of IIC versus CCC for the $A'_{\text{-novel}}$ measure of sensitivity for same exemplars, $F(1, 80) = 6.59$, $MSE = .004$, $p = .01$, with no Experiment \times Condition interaction ($F < 1$).

The combined experiment analyses appear to argue against something akin to a verbal overshadowing account of the findings. When we combined across Experiments 1 and 2, the recognition advantage shown by the III condition over the CCI condition was significant on both the $A'_{\text{-novel}}$ and the $A'_{\text{-related}}$ measures of sensitivity. More important, this item-specific retrieval advantage was unaffected by the alteration in instructions to focus especially on perceptual (shape or visual) information rather than conceptual information (no Condition \times Experiment interaction).

Nonetheless, one might counter that the largely parallel effects under the perceptual versus conceptual instructions provide only slight evidence against something like a verbal overshadowing account. Given the nature of the stimuli—pictures of common objects—one might argue that rapid and spontaneous access to the names of the objects would have made it difficult or impossible to focus only on the shape of the objects in making gist-based recognition decisions (cf. Koutstaal et al., 2003). If, despite the change in instructional emphasis, the names of the objects still tended to guide responding in the category-based condition, then participants in the CCI condition still would show a detrimental effect from earlier having engaged in relatively more verbally mediated prior processing when given the final identical recognition test compared with participants in the III condition, who could have focused on object representations more closely tied to the original pictorial information throughout all three tests.

Experiments 3 and 4 provide a stronger test of the verbal overshadowing account. If the outcomes are attributable to a process similar to that underlying the verbal overshadowing effect (but one that, unlike typical demonstrations of that effect, is induced at the time of retrieval rather than encoding), then—given that verbal overshadowing applies only to complex perceptual or sensory stimuli that are difficult to fully or adequately capture with words—no detrimental effects of the similarity retrieval orientation in the CCI condition compared with the III condition should be observed if the stimuli are themselves words. In contrast, those effects should be at least as great or noticeably exacerbated if participants were asked to overtly provide verbal labels to pictured objects during conceptual recognition. We first examine whether the detrimental effects of sustained conceptual retrieval extend to words rather than pictures (Experiment 3) and then test the effects of requiring overt verbal labeling of pictured objects during conceptual retrieval (Experiment 4).

Experiment 3: Words Rather Than Pictures

The stimulus items in Experiments 1 and 2 shared both perceptual similarity (e.g., all teapots have spouts in approximately similar places) and conceptual, lexical, or semantic similarity (e.g., a shared object name and function). Was perceptual or conceptual similarity more important, or were both necessary to the effects that were observed? Are the detrimental effects of a meaning-based focus confined to basic-level (Rosch, 1978; Rosch et al., 1976) object categories, perhaps because for these types of categories the “meaning comes for free” or very readily and near effortlessly, and additional explicit attention to this information is detrimental (cf. Gosselin & Schyns, 2001; Murphy & Smith, 1982)? Alternatively, do similar detrimental effects also emerge if similarity in meaning derives from shared semantic attributes more generally, and not simply object category membership—in particular for words that are synonyms or near synonyms that have little perceptual similarity across items within a pair?

In Experiment 3 we used synonyms rather than pictures of common objects to provide a key test of the generality of the observed effect on multiple levels (words vs. pictures, basic object level vs. other types of shared meaning). The incidental encoding task was a syllable classification task requiring participants to first read the presented words and then to decide whether the word had more than two syllables. Thereafter, participants were administered three successive episodic recognition tests in the same four between-subjects conditions as in the prior experiments.

One possibility, as we have suggested, is that the detrimental effects of a sustained meaning-based orientation are restricted to instances involving basic-level concepts and pictorial stimuli; in this case, no detrimental effects should be found when the stimuli are synonyms. This is also the prediction that one would make from the verbal overshadowing account. However, another possibility is that similar effects will be observed for stimuli that do not belong to categories in this simple manner. It is even possible that the detailed and individuating perceptual information in pictures might have partially acted to reduce the effects of the meaning-based focus and that even stronger detrimental meaning-focus effects will be observed for synonyms. Either outcome will be highly informative.

Method

Participants. Participants included 72 younger adults (mean age = 20.97 years, $SD = 2.71$) recruited and screened in the same manner as the previous experiments.² They had, on average, 15.93 years of formal education ($SD = 2.25$; data were unavailable for 2 participants).

Stimuli. A total of 240 synonyms and near synonyms (e.g., *brook–stream*, *robber–thief*, *shop–store*), plus practice items, were compiled from various sources. The words were assigned to stimulus sets (for counterbalancing across studied and nonstudied status) and stimulus subsets (for counterbalancing across tests) in the same manner as for the pictorial stimuli in the previous experiments. Stimulus sets and subsets were equated on number of syllables, imageability, and frequency of occurrence, as given in the norms of Kucera and Francis (1967) and G. D. A. Brown (1984) for frequency in written language and listed in the interactive MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm).

Procedure. All general aspects of the procedure were held constant with those of the first two experiments. However, to accommodate the use of the word stimuli, we altered the instructions for the encoding task to a syllable judgment task in which participants were to decide whether the words had more than two syllables. Eighteen participants were tested in each of the four between-subjects conditions (CCC, III, CCI, and IIC).

Results

Effects of prior retrieval orientation on item-specific recognition. Recognition performance with the words was notably reduced compared with that observed for the pictures of common objects used in the two previous experiments (see Figure 3). On the third test, the III and CCI retrieval conditions showed no systematic difference in sensitivity to same exemplars, either relative to new items ($A'_{\text{-novel}}$, III = .73, CCI = .75; $F < 1$) or relative to the related lures ($A'_{\text{-related}}$, III = .75, CCI = .72; $F < 1.2$). However, the CCI group showed significantly greater gist sensitivity on the final test ($A'_{\text{-gist}} = .55$) than did the III group ($A'_{\text{-gist}} = .47$), reflecting higher incorrect endorsements of the related lures for those who had earlier undertaken conceptual compared with identical recognition, $F(1, 34) = 5.50$, $MSE = .01$, $p = .03$, $d = 0.80$. There were no significant effects on any of the bias measures ($F_s < 2.1$).³

Effects of prior retrieval orientation on gist-based recognition. Gist-based recognition of the words was similar to that for the item-specific tests, particularly for same exemplars ($A'_{\text{-novel}}$, CCC = .71, IIC = .71). There was no effect of retrieval condition on this measure ($F < 1$) or on the other sensitivity measures ($A'_{\text{-related}}$, CCC = .63, IIC = .66; $F < 1.3$; $A'_{\text{-gist}}$, CCC = .61, IIC = .57; $F < 1.5$). There were no effects on any of the bias measures ($F_s < 1$).⁴

Given the similar levels of recognition performance for the item-specific and gist-based recognition conditions, one might infer that participants in both conditions were largely responding on the basis of gist. However, positive recognition responses to the different exemplars (synonyms) under the conceptual recognition instructions (where such responses were correct) were significantly more frequent than under the item-specific instructions

² Participants in Experiment 3 were tested in two phases, 12 per condition in the first phase and 6 per condition in the second phase, from a similar participant population. Within each phase, equal numbers of participants were allocated randomly to each of the four conditions.

³ It is necessary to exercise caution in concluding that for synonyms or near synonyms there are no other effects of prior retrieval set on sensitivity beyond the significant gist sensitivity ($A'_{\text{-gist}}$) effect shown here, because, particularly for smaller effects, there is low power to detect a difference (e.g., for a medium effect size of .50, power is approximately .30). On the one hand, for the $A'_{\text{-related}}$ measure, there is a numerical prior retrieval specificity advantage in the comparison of CCI ($A'_{\text{-related}} = .72$) versus III ($A'_{\text{-related}} = .75$), paralleling the significant specificity advantage found for this comparison with objects in Experiments 1 and 2. On the other hand, for the $A'_{\text{-novel}}$ measure, the direction of the means slightly favors prior conceptual retrieval.

⁴ The pattern of sensitivity observed on the final conceptual test for the CCC versus IIC groups is noteworthy: Whereas participants who had previously focused on item-specific information continued to show numerically greater sensitivity to the difference between synonyms on the conceptual test ($A'_{\text{-related}} = .66$) than did those who had focused on conceptual information throughout ($A'_{\text{-related}} = .63$), the reverse pattern was observed for those who had previously focused on conceptual information: These participants showed numerically greater sensitivity to gist on the final test ($A'_{\text{-gist}} = .61$) than did participants who initially focused on item-specific information ($A'_{\text{-gist}} = .57$). It is possible that with a larger sample size these differences in sensitivity would prove significant. Should such differences be observed, they (like the significant results for the $A'_{\text{-gist}}$ measure for the III vs. CCI comparison we have reported) would further argue against the feasibility of a verbal overshadowing account of the results.

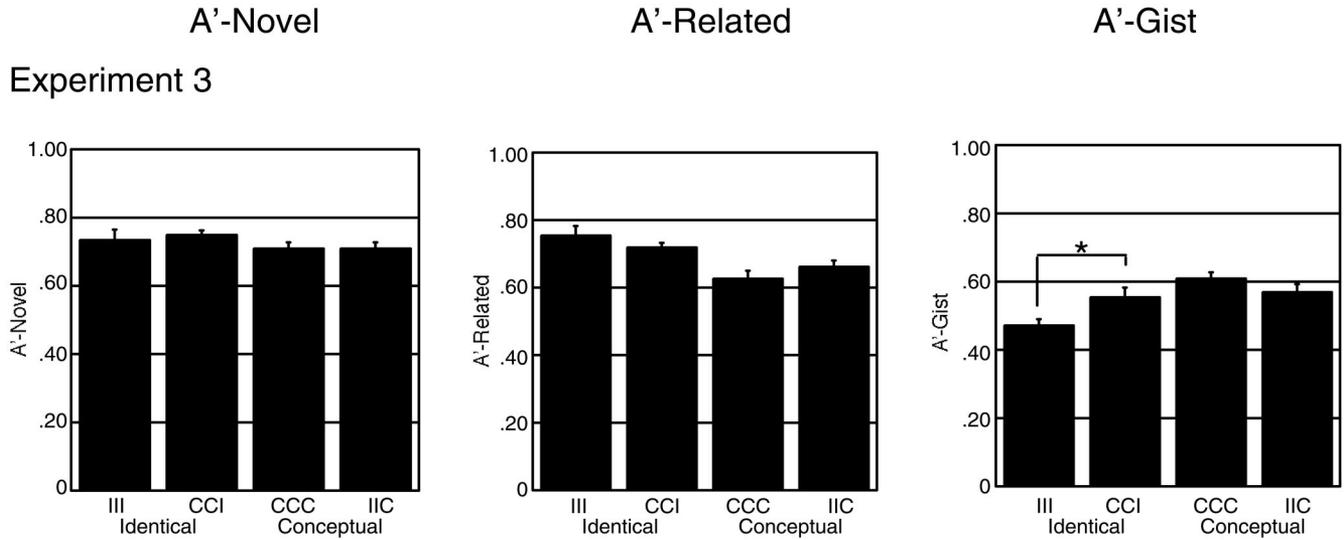


Figure 3. Measures of sensitivity ($A'_{\text{-novel}}$, $A'_{\text{-related}}$, and $A'_{\text{-gist}}$) for Experiment 3, which used word stimuli. Sensitivity measures are shown separately for each of the four conditions, in which the type of recognition probe either remained constant (III = identical, identical, identical; CCC = conceptual, conceptual, conceptual) or changed on the final test (IIC = identical, identical, conceptual; CCI = conceptual, conceptual, identical). The measures of sensitivity shown are for the final (third) test in each condition; the prior two tests established an orientation to either item-specific or conceptual information. A' of .50 represents chance performance. Error bars show the standard error of the mean. $*p < .05$.

(where such responses constituted false recognitions), $F(1, 70) = 11.18$, $MSE = .03$, $p = .001$, $d = 0.80$. These outcomes indicate that participants were, indeed, able to perform the two recognition tasks for the word stimuli and that they were able to differentially call on item-specific information when needed to reject successfully at least some similar-appearing lures.

Discussion

Unlike the patterns found in Experiments 1 and 2, for the measures of sensitivity to item-specific information ($A'_{\text{-novel}}$ and $A'_{\text{-related}}$), there was no evidence of carryover effects from a prolonged retrieval orientation toward either item-specific or conceptual information when the stimuli were synonyms or near synonyms. However, in a more indirect parallel to the item-specific retrieval advantage shown in Experiments 1 and 2, there was a significant advantage for prior item-specific retrieval on the measure of sensitivity to gist on the final identical test, with the CCI group showing greater carryover (and now inappropriate) gist sensitivity than the III group.

These findings suggest that it may be difficult to recover from the adoption of a gist-based retrieval set for verbal materials if it subsequently proves necessary to make more fine-grained item-specific recognition decisions among semantically similar alternatives. The emergence of a significant detrimental carryover effect of prior category-based retrieval for the synonym and near-synonym word stimuli used here also points to some generality of the item-specific retrieval advantage (albeit, in this case, as shown by continued but now inappropriate sensitivity to gist rather than decreased sensitivity to item-specific information). Experiment 3 demonstrates that prolonged retrieval at a meaning-based level can be harmful for later item-specific episodic memory not only for stimuli that share basic-level names but also for stimuli for which the similarity in meaning

derives from shared semantic attributes more generally. Equally important, the findings seem to further argue against the viability of a version of the verbal overshadowing account as an explanation of the retrieval specificity advantage. There was again a retrieval specificity advantage (albeit now on the $A'_{\text{-gist}}$ measure for the final item-specific test), even though the stimuli were themselves now words.

However, there were also several other differences between the experiments that used pictures versus words that plausibly might have affected the outcomes. For example, because it was not possible to find 240 pairs of synonyms that denoted concrete nouns, the encoding task for the word stimuli (syllable judgment) differed from that for the picture stimuli (judgments of the size of the object referent). We could provide a cleaner test of the verbal overshadowing account by requiring verbal labeling of the objects at the time of retrieval. According to the verbal overshadowing account, this should heighten the detrimental effect of sustained meaning-based retrieval. The final experiment tests this possibility.

Experiment 4: Conceptual Focus With Verbal Labeling

Method

Participants. The participants were 96 younger adults (mean age = 19.41 years, $SD = 1.39$), recruited and screened in the same manner as in each of the previous experiments. They had, on average, 14.06 ($SD = 1.63$) years of formal education.

Stimuli. The stimuli were the same as those used in Experiments 1 and 2.

Procedure. With one notable exception, the procedure was identical to that used in Experiment 1. The exception concerned instructions to participants for the conceptual recognition tests. Participants who performed these tests were asked to name each of the objects that appeared on the test out loud before making their recognition decisions. Twenty-four participants were tested in each of the four between-subjects conditions (CCC, III, CCI, and IIC).

Results

Effects of prior retrieval orientation on item-specific recognition. Unlike the pattern observed in Experiments 1 and 2, the item-specific recognition performance of participants who engaged in item-specific retrieval throughout the session did not exceed the performance of those who first did conceptually based recognition and then switched to item-specific retrieval on the final test. Indeed, the numerical values for $A'_{\text{-novel}}$ (III = .89, CCI = .91) and $A'_{\text{-related}}$ (III = .83, CCI = .86) were in the opposite direction (see Figure 2), although not significantly so: The largest F ratio was that for $A'_{\text{-novel}}$, $F(1, 46) = 2.65$, $MSE = .002$, $p = .11$, $d = 0.48$, although note that observed power was only .50. The two groups did not differ in gist sensitivity (.68 for both groups) or on any bias measures ($F_s < 1$).

Effects of prior retrieval orientation on gist-based recognition. Again, unlike the pattern observed in Experiments 1 and 2, the conceptually based recognition performance of participants who engaged in item-specific retrieval up until the time of the final test did not exceed the performance of those who engaged in conceptually based recognition throughout. In addition, for the same exemplars, the means again numerically favored the conceptually based retrieval orientation for the $A'_{\text{-novel}}$ measure (CCC = .90, IIC = .88; $F < 1$), although not for the $A'_{\text{-related}}$ measure (CCC = .54, IIC = .55; $F < 1$). For the different exemplars, this nonsignificant reversal was again observed (for $A'_{\text{-gist}}$, CCC = .89, IIC = .87; $F < 1.5$). There were no effects on measures of response bias ($F_s < 1.8$).

Discussion

Contrary to the findings expected on the verbal overshadowing account, the requirement to verbally label the objects during the conceptual recognition tests attenuated or even slightly reversed—rather than accentuating—the prior item-specific retrieval advantage for the III condition compared with the CCI condition. The CCI condition now showed little systematic difference from the III condition. Consistent with this observation, across-experiment analyses considering all three object experiments (Experiments 1, 2, and 4) showed a significant Experiment \times Condition interaction both for the $A'_{\text{-novel}}$ measure, $F(2, 126) = 4.39$, $MSE = .002$, $p = .01$, and the $A'_{\text{-related}}$ measure, $F(2, 126) = 5.24$, $MSE = .005$, $p = .007$. Follow-up comparisons that separately contrasted Experiment 4 with Experiments 1 and 2 showed significant decrements on these sensitivity measures following conceptually focused retrieval only in the earlier experiments and not in Experiment 4 (significant Experiment \times Condition interactions for both comparisons for both measures; smallest $F(1, 92) = 6.18$, $p = .01$).

An important broader inference to be drawn from the pattern observed in Experiment 4 versus Experiment 1 is that individuals likely were not typically uniformly naming items during testing under the conceptual instructions in Experiment 1. If they had been, the requirement to perform naming during retrieval in the current experiment should not have led to such a different pattern of performance from that earlier experiment.

An apparent puzzle is that naming did not impede but rather actually seemed slightly to help subsequent performance on the item-specific task in the CCI condition compared with the III condition; as we have noted, this is directly counter to something akin to a verbal overshadowing account both of our other findings

and of the present findings. However, an important difference between the current paradigm and verbal overshadowing paradigms is that in the present paradigm individuals were asked to engage in labeling at the time of retrieval—sometime after they had already encoded items—not at the time of encoding; encoding was held constant for all of the conditions. It is possible that labeling in this retrieval situation was helpful because it served to introduce new information that, although it did not substantially change the accessibility of previously encoded perceptual information, helped to isolate the relevant items from the populous perceptual world of more than 200 pictorial images that participants were presented across the various recognition tests. That is, in contrast to situations in which labeling or verbal description occurs during initial encoding—and when verbal codes may preempt or interfere with the encoding of complex (not readily verbalized) aspects of stimuli—the emphasis on labels during retrieval in the current paradigm might have provided an additional retrieval cue to the already encoded perceptual information, serving to help participants access that information.

Some support for this suggestion is offered by a comparison of the across-test decrements in performance for the conceptual tests in Experiment 1 versus Experiment 4. Whereas there were clear decrements across successive tests for both same and different exemplars and for both CCC and CCI conditions in Experiment 1, there was perhaps a relative resistance to such decrements in Experiment 4. The average conceptual recognition performance values (combining across same and different exemplars and subtracting false alarms to novel items) for the CCC condition in Experiment 1 were .71, .61, and .53 for Tests 1 through 3, respectively; the corresponding values for the CCC condition in Experiment 4 were .76, .74, and .67, respectively. An across-experiment comparison treating experiment (Experiment 1 vs. 4) as a between-subjects factor and test (Test 1, 2, or 3) as a within-subject factor provided evidence of greater resistance to forgetting or interference under the overt verbal labeling gist-based retrieval conditions of Experiment 4, $F(2, 80) = 4.45$, $MSE = .006$, $p = .01$, for the Experiment \times Test interaction.

General Discussion

The primary purpose of the experiments reported here was to examine the impact of an explicitly and intentionally adopted gist-based versus item-specific retrieval orientation on later knowledge availability. With regard to this aim, the multiple-test episodic recognition paradigm developed and used in these experiments has yielded four main conclusions.

1. Prior retrieval at an item-specific level led to higher item-specific recognition of subsequently probed (and not previously tested) items than did prior retrieval under either general conceptual gist instructions (Experiment 1) or perceptually focused gist instructions (Experiment 2). These effects were especially strong on the measure of sensitivity contrasting hits to same versus related exemplars (a particularly stringent measure of sensitivity to item-specific information) but also were apparent on the measure of sensitivity contrasting hits to same exemplars versus novel items. Analyses that combined the results across Experiments 1 and 2 showed a significant item-specific retrieval advantage on both measures, with no interactions.

2. Prior item-specific retrieval also benefited gist-based retrieval such that gist-based retrieval on the final critical test was higher if

earlier testing of episodically related items took place under item-specific rather than gist instructions. These effects were especially observed on the $A'_{\text{-novel}}$ measure, with this measure showing a significant prior item-specific retrieval advantage in Experiment 1, a numerical advantage in Experiment 2, and a significant advantage in combined analyses of these two experiments.

3. However, item-specific retrieval no longer carried this edge when the gist-based task required participants to overtly name each object directly before making their recognition decisions (Experiment 4). Under these conditions there was, instead, a modest (nonsignificant) advantage to the gist-based retrieval approach for both the item-specific and the gist-based final test.

4. When words (synonym pairs) rather than objects (paired exemplars) were used, there were no systematic effects of prior retrieval orientation on the two measures of item-specific memory that showed effects of prior retrieval orientation for pictured objects ($A'_{\text{-novel}}$ and $A'_{\text{-related}}$). Nonetheless, there was again evidence for an advantage of prior item-specific retrieval on the $A'_{\text{-gist}}$ measure. Participants who sustained an item-specific retrieval orientation throughout (III condition) made significantly fewer gist-based errors on the final item-specific test than did participants who initially engaged in category-based retrieval and then attempted to switch to a more specific retrieval focus on the final test (CCI condition).

A more general observation that we might make on the basis of the findings for Experiments 1, 2, and 4, each of which used gist-based instructions for pictured objects, concerns the specificity of the gist-based instructions themselves. Differential emphasis within the instructions and different task requirements (e.g., the requirement to overtly name objects) led to noticeably different patterns of performance. A comparison with incidental encoding instructions may be instructive here. Adoption of a specific encoding task helps to ensure relatively constant encoding of particular attributes during the study phase (Craik & Lockhart, 1972; Lockhart & Craik, 1990). Similar considerations may apply to instructions at retrieval. When the encoded stimuli are complex and multidimensional, instructions to make meaning-based, category-based, or gist-based retrieval decisions may be relatively underspecified, leaving considerable leeway for person-to-person and item-to-item variation in the relative weighting of a given dimension or attribute in the recognition classification decision.

One of our primary conclusions from these experiments is very similar to that drawn by Whittlesea et al. (1994) in a rare exploration (with a very different paradigm) of the possible conjoint application of both general and particular knowledge during retrieval. These researchers argued that “no simple correspondence of knowledge types and demands to produce class-level or item-level judgments will capture the variety of interactions between previously acquired knowledge and the cues that control its application,” further urging that “careful attention to details of the processing situation is potentially very important in understanding the use of everyday knowledge” (p. 260; also see Barnard, 1985, 1998; Whittlesea & Leboe, 2000). Even probing recognition memory of healthy young adults for highly memorable pictorial stimuli drawn from well-known semantic categories, tested within a single experimental session, showed alterations in ease of access to knowledge depending on the specificity of the recognition tests for items from the episode and details of the probe itself (e.g., differential emphasis on category names vs. object-level representations). Moreover, the level of grain at which participants had

previously probed their memory for an episode had carryover effects for subsequent attempts to retrieve other—not previously retrieved—items from the episode.

Prior retrieval at a coarser, more categorical level of specificity served to decrease the likelihood of subsequent item-specific recognition of items that had not been previously tested, both when the prior retrieval was guided toward general conceptual similarities (Experiment 1) and when it was guided toward similarities in object shape or form (Experiment 2). Further, it was not simply that initial item-specific retrieval facilitated subsequent item-specific retrieval; the prior item-specific retrieval orientation advantage also extended, albeit less strongly and consistently, to the gist-based recognition tests, such that participants who initially had engaged in item-specific retrieval showed higher correct conceptual recognition on the final test than did those who consistently engaged in gist-based retrieval. This item-specific retrieval advantage was significant for the general conceptual similarity instructions of Experiment 1; it was also significant, for the $A'_{\text{-novel}}$ measure, when we combined across Experiments 1 and 2.

These demonstrations of a prior retrieval orientation effect on recognition memory for other episodically related items extend the findings of the many studies that have shown that there may be detrimental effects of prior retrieval of some items on the subsequent retrieval chances of other items that were not previously tested (e.g., M. C. Anderson et al., 1994; Koutstaal, Schacter, Johnson, & Galluccio, 1999; Perfect et al., 2004) and show that it is not only the mere occurrence or the frequency of prior retrieval efforts but also the specificity at which such retrieval occurs that is important.

A somewhat related precedent for this conclusion, but in the domain of free recall, was reported by Whitten and Leonard (1980). They showed greater improvement on an unexpected free recall test if participants had been tested earlier on a multiple-choice recognition test that required particular care in selecting the correct response because of the presence of a larger number of semantically related alternatives. However, in that study, the authors’ conclusion that “long-term retention is facilitated to the extent that memory evaluation occurs during initial recognition tests” (p. 127) was based on prior tests of the target items themselves rather than items that had never been previously tested but were episodically related.

In some respects, the current outcomes may cohere with other conclusions that have emerged from explorations of the verbal overshadowing phenomenon, that changes in retrieval orientation may have more global effects, including transfer-inappropriate retrieval effects (e.g., Schooler, 2002). However, the entire set of results we report cannot simply be explained from this perspective. One difficulty for such an account is the evidence of an item-specific retrieval advantage for conceptual recognition (especially in Experiment 1). It seems that prior conceptual retrieval should have more positive retrieval overlap for subsequent conceptual retrieval, but in Experiments 1 and 2 prior item-specific retrieval yielded higher levels of conceptual retrieval on the final critical test. Another clear difficulty for this account is that the increased verbal retrieval emphasis in Experiment 4 (conceptual recognition with concurrent overt labeling) did not impede but rather seemed to enhance recognition.

The concurrent overt labeling manipulation in Experiment 4 might have led to a more uniform retrieval set compared with the more generic and less constrained conceptual (Experiment 1) or

perceptual (Experiment 2) gist instructional sets. There also was evidence that the emphasis on overt labeling might have reduced across-test interference by introducing new cuing information that helped to isolate the relevant items from the many perceptual images presented across the entire testing session. Whatever the origins of the slight conceptual recognition performance enhancement found under the overt naming condition, the contrasting pattern in this condition relative to the conceptual retrieval conditions in Experiments 1 and 2 suggests that, left to their own devices, younger participants given gist-based retrieval instructions for pictorial items with known semantic content likely do not always use verbal labels to mediate their performance on the task.

Taken together, the outcomes from all of the experiments reported here argue that a full understanding of the ways prior retrieval may affect subsequent memory—even for events or facts that were not themselves retrieved—cannot only focus on the mere “fact” of retrieval (whether retrieval occurred) but must also take into account the *specificity* with which that retrieval occurred. Although, depending on the tasks at hand, there may be important reasons to encourage gist-based retrieval, that very focus may itself be detrimental for other tasks in which item-specific information might be crucial.

From this perspective, more recent emphases on the need to focus on both more specific and more abstract knowledge if one hopes to maximally encourage the development and transfer of expertise may already represent a key theoretical and practical acknowledgment of the importance of this constraint. Interleaved exposure to abstract (rule- or heuristic-based) and specific (instance-based) training, or *strategically filtered* exposure to selective or reduced-context specific instances, has been found to yield both more efficient learning and more appropriate use of knowledge than does an exclusive focus on either level alone (e.g., Effken, Kim, & Shaw, 1997; Kostopoulou & Duncan, 2001; Munley & Patrick, 1997; also cf. Beishuizen, Asscher, Prinsen, & Elshout-Mohr, 2003).

At a more applied level, these outcomes have an important implication that warrants systematic evaluation: Cognitive, neuropsychological, or other factors that initially lead to the adoption of high and sustained levels of gist-based processing (e.g., clinical depression, cognitive aging) may cumulatively make it more difficult to access specific information from memory for reasons relating to alterations in retrieval processes such as those underlying the current outcomes, independent of (or in addition to) the original factors that led to greater gist-based processing. For instance, in clinical depression, a sustained orientation toward the retrieval of general or categorical memories (Williams & Dritschel, 1992) may make it increasingly difficult to retrieve specific episodes or details, further compounding the impairments in problem solving that overly general memory retrieval may yield (e.g., Raes et al., 2005; Sidley, Whitaker, Calam, & Wells, 1997; see, especially, Watkins, Teasdale, & Williams, 2000, for evidence that alterations of cognitive state at the time of memory retrieval may reduce the recall of categorical memories in depression). Likewise, a gist-based predisposition in older adults (e.g., Koutstaal et al., 2003; Tun et al., 1998) may, across time, make more specific retrieval ever more difficult.

The need for meaning-based abstraction is undeniable in a wide array of tasks, ranging from categorical inference to analogy use to metaphorical interpretation (e.g., Brainerd & Reyna, 1990, 1992). Yet the central role of specific information and instances in these

and other forms of thinking is likewise increasingly recognized (e.g., Goldstone & Barsalou, 1998). Given these contrary poles and apparently conflicting processing goals, it may be that what is most needed for optimal thinking and problem solving, and possibly also optimal memory retention, is not one or the other but rather *flexible remembering* (Koutstaal, 2006): an ability to adeptly and rapidly move between specific and abstract memory retrieval as needed. Indeed, recent findings with older adults have shown that flexible remembering in response to changing task demands or goals is significantly positively correlated with measures of semantic and phonological fluency (Koutstaal, 2006)—themselves indexes of the ease with which people can strategically search for and flexibly bring to mind well-learned information in accord with particular contextual requirements (Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Perhaps with a more equal weighting and valuing of both specificity and abstraction, we will have a greater chance to understand how it is that we do (and sometimes, still, do not) use what we know.

References

- Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 8, 463–470.
- Anderson, M. C., & Bell, T. (2001). Forgetting our facts: The role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General*, 130, 544–570.
- Anderson, M. C., Bjork, R. R., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1063–1087.
- Anderson, S. J., & Conway, M. A. (1993). Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1178–1196.
- Barnard, P. (1985). Interacting cognitive subsystems: A psycholinguistic approach to short-term memory. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp. 197–258). Hove, England: Erlbaum.
- Barnard, P. J. (1998). Interacting cognitive subsystems: Modelling working memory phenomena within a multi-processor architecture. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 298–339). Cambridge, England: Cambridge University Press.
- Bäumli, K. H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*, 13, 356–360.
- Beishuizen, J., Asscher, J., Prinsen, F., & Elshout-Mohr, M. (2003). Presence and place of main ideas and examples in study texts. *British Journal of Educational Psychology*, 73, 291–316.
- Brainerd, C. J., & Gordon, L. L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology*, 30, 163–177.
- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the gist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, 10, 3–47.
- Brainerd, C. J., & Reyna, V. F. (1992). Explaining “memory free” reasoning. *Psychological Science*, 3, 332–339.
- Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review*, 100, 42–67.
- Brainerd, C. J., & Reyna, V. F. (1998). When things that were never experienced are easier to “remember” than things that were. *Psychological Science*, 9, 484–489.
- Brown, C., & Lloyd-Jones, T. J. (2002). Verbal overshadowing in a multiple face presentation paradigm: Effects of description instruction. *Applied Cognitive Psychology*, 16, 873–885.
- Brown, G. D. A. (1984). A frequency count of 190,000 words in the

- London-Lund Corpus of English Conversation. *Behavioral Research Methods Instrumentation and Computers*, 16, 502–532.
- Brown, J. (1968). Reciprocal facilitation and impairment of free recall. *Psychonomic Science*, 10, 41–42.
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1147–1156.
- Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1403–1414.
- Cohen, G. (2000). Hierarchical models in cognition: Do they have psychological reality? *European Journal of Cognitive Psychology*, 12, 1–36.
- Conway, M. A. (1992). A structural model of autobiographical memory. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 167–193). Dordrecht, the Netherlands: Kluwer Academic.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13, 595–605.
- Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, 25, 129–139.
- Dorfman, J., & Mandler, G. (1994). Implicit and explicit forgetting: When is gist remembered? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 47(A), 651–672.
- Effken, J. A., Kim, N. G., & Shaw, R. E. (1997). Making the constraints visible: Testing the ecological approach to interface design. *Ergonomics*, 40, 1–27.
- Ewa, K. W. (2002). The aging physician: Changes in cognitive processing and their impact on medical practice. *Academic Medicine*, 77(Suppl. S), S1–S6.
- Fallshore, M., & Schooler, J. W. (1995). The verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1608–1623.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General*, 131, 73–95.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65, 231–262.
- Gosselin, F., & Schyns, P. G. (2001). Why do we SLIP to the basic level? Computational constraints and their implementation. *Psychological Review*, 108, 735–758.
- Hogan, R. M., & Kintsch, W. (1971). Differential effects of study and test trials on long-term retention and recall. *Journal of Verbal Learning and Verbal Behavior*, 10, 562–567.
- Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, 126, 248–277.
- Karchmer, M. A., & Winograd, E. (1971). Effects of studying a subset of familiar items on recall of the remaining items: The John Brown effect. *Psychonomic Science*, 25, 224–225.
- Koriat, A., & Goldsmith, M. (1996). Memory metaphors and the real life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences*, 19, 167–188.
- Kostopoulou, O., & Duncan, K. D. (2001). Abstract and reduced-context representations in fault-finding training. *Ergonomics*, 44, 175–201.
- Koutstaal, W. (2003). Older adults encode—but do not always use—perceptual details: Intentional versus unintentional effects of detail on memory judgments. *Psychological Science*, 14, 189–193.
- Koutstaal, W. (2006). Flexible remembering. *Psychonomic Bulletin & Review*, 13, 84–91.
- Koutstaal, W., Reddy, C., Jackson, E. M., Prince, S., Cendan, D. L., & Schacter, D. L. (2003). False recognition of abstract versus common objects in older and younger adults: Testing the semantic categorization account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 499–510.
- Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, 37, 555–583.
- Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer, K. A. (1999). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging*, 14, 220–237.
- Koutstaal, W., Schacter, D. L., Johnson, M. K., Angell, K. E., & Gross, M. S. (1998). Post-event review in older and younger adults: Improving memory accessibility of complex everyday events. *Psychology and Aging*, 13, 277–296.
- Koutstaal, W., Schacter, D. L., Johnson, M. K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition*, 27, 478–493.
- Koutstaal, W., Verfaellie, M., & Schacter, D. L. (2001). Recognizing identical versus similar categorically related common objects: Further evidence for degraded gist-representations in amnesia. *Neuropsychology*, 15, 268–289.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kulatunga-Moruzi, C., Brooks, L. R., & Norman, G. T. (2001). Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teaching and Learning in Medicine*, 13, 110–116.
- Linton, M. (1975). Memory for real-world events. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 376–404). San Francisco: Freeman.
- Linton, M. (1978). Real world memories after six years: An in vivo study of very long-term memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). New York: Academic Press.
- Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology*, 44, 87–112.
- Macrae, C. N., & Lewis, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science*, 13, 194–196.
- Matthews, R. C. (1991). The forgetting algorithm: How fragmentary knowledge of exemplars can abstract knowledge. *Journal of Experimental Psychology: General*, 120, 117–119.
- Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 189–223). Hillsdale, NJ: Erlbaum.
- Micco, A., & Masson, M. E. J. (1992). Age-related differences in the specificity of verbal encoding. *Memory & Cognition*, 20, 244–253.
- Mueller, C. W., & Watkins, M. J. (1977). Inhibition from part-set cuing: A cue-overload interpretation. *Journal of Verbal Learning and Verbal Behavior*, 16, 699–709.
- Munley, G. A., & Patrick, J. (1997). Training and transfer of a structural fault-finding strategy. *Ergonomics*, 40, 92–109.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Murphy, G. L., & Smith, E. E. (1982). Basic level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior*, 21, 1–20.
- Pansky, A., & Koriat, A. (2004). The basic-level convergence effect in memory distortions. *Psychological Science*, 15, 52–59.
- Perfect, T. J., Stark, L.-J., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent

- nature of retrieval-induced forgetting. *Journal of Memory and Language*, 51, 399–417.
- Raes, F., Hermans, D., Williams, J. M. G., Demyttenaere, K., Sabbe, B., Pieters, G., & Eelen, P. (2005). Reduced specificity of autobiographical memory: A mediator between rumination and ineffective social problem-solving in major depression? *Journal of Affective Disorders*, 87, 331–335.
- Ramponi, C., Barnard, P. J., & Nimmo-Smith, I. (2004). Recollection deficits in dysphoric mood: An effect of schematic models and executive mode? *Memory*, 12, 655–670.
- Reder, L. M., Wible, C., & Martin, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 72–81.
- Reyna, V. F. (2004). How people make decisions that involve risk: A dual-process approach. *Current Directions in Psychological Science*, 13, 60–66.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7, 1–75.
- Roediger, H. L., III, Jacoby, J. D., & McDermott, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language*, 35, 300–318.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, 16, 989–997.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36–71.
- Schooler, J. W., Foster, R. A., & Loftus, E. F. (1988). Some deleterious consequences of the act of recollection. *Memory & Cognition*, 16, 243–251.
- Schreiber, T. A., & Sergent, S. D. (1998). The role of commitment in producing misinformation effects in eyewitness memory. *Psychonomic Bulletin & Review*, 5, 443–448.
- Shaw, J. S., III, Bjork, R. A., & Handal, A. (1995). Retrieval induced forgetting in an eyewitness memory paradigm. *Psychonomic Bulletin & Review*, 2, 249–253.
- Sidley, G. L., Whitaker, K., Calam, R. M., & Wells, A. (1997). The relationship between problem-solving and autobiographical memory in parasuicide patients. *Behavioural and Cognitive Psychotherapy*, 25, 195–202.
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457–482.
- Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, 36, 499–504.
- Tuckey, M. R., & Brewer, N. (2003). How schemas affect eyewitness memory over repeated retrieval attempts. *Applied Cognitive Psychology*, 17, 785–800.
- Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, 13, 230–241.
- Watkins, E., Teasdale, J. D., & Williams, R. M. (2000). Decentering and distraction reduce overgeneral autobiographical memory in depression. *Psychological Medicine*, 30, 911–920.
- Westerman, D. L., & Larsen, J. D. (1997). Verbal-overshadowing effect: Evidence for a general shift in processing. *American Journal of Psychology*, 110, 117–128.
- Whitten, W. B., & Leonard, J. M. (1980). Learning from tests: Facilitation of delayed recall by initial recognition alternatives. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 127–134.
- Whittlesea, B. W. A., Brooks, L. R., & Westcott, C. (1994). After the learning is over: Factors controlling the selective application of general and particular knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 259–274.
- Whittlesea, B. W. A., & Leboe, J. P. (2000). The heuristic basis of remembering and classification: Fluency, generation, and resemblance. *Journal of Experimental Psychology: General*, 129, 84–106.
- Williams, J. M. G., & Broadbent, K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, 95, 144–149.
- Williams, J. M. G., & Dritschel, B. H. (1992). Categorical and extended autobiographical memories. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 391–410). Dordrecht, the Netherlands: Kluwer Academic.
- Yaniv, L., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy–informativeness trade-off. *Journal of Experimental Psychology: General*, 124, 424–432.

Received November 10, 2004

Revision received September 20, 2005

Accepted January 3, 2006 ■