ABSTRACT—We examine whether temporally defined associations play a role in item recognition. The role of these associations in recall tasks is well known; we demonstrate an important role in item recognition as well. In this study, subjects were significantly more likely to recognize a test item as having been previously experienced if the preceding test item was studied in a temporally proximal list position than if the preceding test item came from a more distant list position. Further analyses showed that this associative effect was almost entirely due to cases in which the preceding test item received a highest-confidence recognition judgment.

In old/new recognition, a subject studies a series of items and is then shown a series of test probes. Subjects judge whether each probe is "old" (a repetition of an item that was presented on the studied list) or "new" (an item that had not appeared during the experiment). Widely studied for almost 100 years (Strong, 1912), the old/new recognition task is believed to measure item-specific memory, devoid of interitem associations (Humphreys, 1978; Murdock, 1974). In contrast, recall tasks rely on strong temporal associations among items (Raskin & Cook, 1937). For example, in the free-recall task, recall of a list item tends to be followed by recall of an item studied in a nearby list position, even though the instructions permit subjects to recall items in any order they wish (Kahana, 1996).

Many models of recognition and recall capture this distinction between item and associative information (Gillund & Shiffrin, 1984; Hintzman, 1983; Humphreys, Pike, Bain, & Tehan, 1989; Murdock, 1982; Shiffrin & Steyvers, 1997). In these models, item information reflects a weighted summed similarity between the probe item and the contents of memory. Associative information is typically stored independently by means of a conjunctive process that binds the information making up the individual items. Associations in these models not only are formed between simultaneously processed items, but also can span several items that co-occur in a working memory buffer (e.g., Kahana, 1996; Siroton, Kimball, & Kahana, in press).

Although many models assume a strict boundary between item information, driving recognition, and associative information, driving recall, it has long been recognized that recognition and recall are complex tasks that reflect the interaction of multiple memory systems, processes, or operations (see Kahana, Rizzuto, & Schneider, in press, for a review). In the case of recognition memory, a major line of research assumes that two distinct item-specific processes drive recognition judgments: a fast acting familiarity process and a slower recollective process (Arndt & Reder, 2002; Atkinson & Juola, 1974; Mandler, 1980; Yonelinas, 2001). Familiarity in these models coarsely maps onto the summed-similarity notion described in the preceding paragraph. Recollection is a recall-like process that involves the recovery of specific source information about the remembered item and is accompanied by a conscious experience of having seen or heard the target item (Tulving, 1985).

What exactly is the nature of the recollective process that underlies some recognition judgments? We consider here the possibility that recollective experience reflects, in part, the shared information between items studied in temporal proximity rather than simply the detailed features of an individual item's occurrence. In recall tasks, such associative features may be seen in the strong tendency for recall of an item to evoke memories of its neighbors (Kahana, 1996). To the extent that recognition of an item as "old" involves a recall-like recollection process, one might observe temporally defined retrieval effects in item recognition, despite the fact that the task does not explicitly test associative information. This could point to the existence of common associative mechanisms in associative recall and item recognition. Thus, our main goal in this article is to examine the effects of temporal co-occurrence on recognition performance.

Light and Schurr (1973) offered some evidence for the possible role of temporal associative mechanisms in an item
recognition task. For one group of subjects, old test probes were presented in the same order as they were studied; for a second group, the order of the old test probes was randomized. Subjects in the same-order group performed better on the recognition test than subjects in the random-order group. In a related study, Ratcliff and McKoon (1978) had subjects make recognition judgments on nouns read as part of a story. They found that recognition memory for an old item was enhanced when the previous old-item probe came from within the same linguistic unit (i.e., from the same proposition or sentence rather than a different proposition or sentence). With these linguistically organized materials, Ratcliff and McKoon did not find an effect of within-proposition interword distance. Their effect, rather, was carried by the co-occurrence of items within a given proposition or sentence.

To assess the effect of temporal co-occurrence on recognition performance, we conducted a recognition memory experiment using pictures as stimuli. The recognition test was a random sequence of test probes that included the old items from the list intermingled with an equal number of new items that served as lures. Subjects pressed one of six keys in response to each probe, rating their confidence that they had seen it before from 1 (sure new) to 6 (sure old). The subjects’ ability to discriminate the old pictures from the new pictures had to be a consequence of their memory for the studied items. A recognition test might include the subsequence of test probes (⋯, O₁₂, N, O₁₇, N, N, O₉₉, ⋯), where N denotes a new item and Oᵧ denotes an old item from position 𝑥 in the study list. We define the lag, r, between two successive old items (⋯, Oᵢ, Oⱼ, ⋯) as the distance, 𝑗−𝑖, between the items on their initial presentation (see Fig. 1). Suppose that recognition of a test item, Oᵢ, brings forth the mental state that prevailed when Oᵢ was first encoded. Suppose further that this retrieved mental state contributes to the retrieval environment that determines subsequent recognition judgments. Then,

![Fig. 1. Illustration of the item recognition task and the calculation of lag. After studying a series of pictures, subjects judged each picture in a test series as “old” (previously studied) or “new” (not previously studied). We define lag as the distance in the study list between successive “old” test items.](image-url)
if the very next test item is $O_j$, we would predict that memory for $O_j$ should be enhanced when $r$ is near zero.

**EXPERIMENT**

**Method**

**Subjects**

Ninety-one Syracuse University undergraduates participated for course credit. The study was approved by the Syracuse University Institutional Review Board. All subjects provided informed consent.

**Procedure**

Each subject studied six lists of 64 digital pixmaps with $350 \times 232$ resolution. The images subtended approximately 5° of visual angle. Images were obtained from planetware.com, a travel-picture Web site, and presented on the screen for 1 s, with a 0.5-s blank interstimulus interval. A sequence of 128 test probes were given immediately following each study list. Half of the test probes were previously studied, old, items; the remainder were new.

A computer algorithm constructed the test lists separately for each subject. Old/new status was first randomly assigned to each test position. Old items were then assigned to test positions in two stages. First, randomly chosen pairs of old items with lags from $-5$ through $5$ were placed in adjacent test positions that had been previously designated for old items. Second, the remaining test positions that were designated to have old items were filled with the remaining old items, and the test positions designated to have new items were filled with new pictures. This algorithm ensured that we would have sufficient data for each subject to assess the effect of lag within the range of $-5$ through $5$.

In response to each test picture, subjects pressed a button from “1” to “6” to describe their confidence that the test item had been previously studied (sure old) and a response of “1” when they were absolutely certain that the test picture was not previously studied (sure new). The keys were arranged to allow subjects to respond comfortably by using the first three fingers of each hand. Responses with latencies longer than 3 s or shorter than 50 ms were discarded. Subjects studied and were tested on a practice list prior to the first study list. They were instructed to use all six buttons and to respond as quickly as possible without sacrificing accuracy. After the practice session, subjects were given feedback concerning their mean response time (RT) and the distribution of their responses across confidence levels.

**Results**

Subjects’ mean hit rate (HR; probability of correctly responding “4,” “5,” or “6” to a new item) was .68. Subjects’ mean false alarm rate (FAR; probability of incorrectly responding “4,” “5,” or “6” to a new item) was .28. The mean value of $d’$ was 1.11. Mean RTs for hits, misses, false alarms, and correct rejections were 1.05, 1.17, 1.24, and 1.17 s, respectively. RTs were significantly shorter for hits than for all other response types; RTs were significantly longer for false alarms than for all other response types, $t(90) > 4.5$, $p < .001$. RTs for misses and correct rejections did not differ significantly ($t = 0.60$).

To assess the effect of temporal co-occurrence on recognition performance, we compared responses to old items following items from adjacent list positions ($l|l = 1$) with responses to old items following items from remote positions ($l|l > 10$). Figure 2 shows receiver-operating-characteristic (ROC) functions for these two classes of old items. ROC functions describe HR as a function of FAR separately for each confidence level. That is, for each confidence level, we counted all confidence judgments greater than or equal to that confidence level as “old” responses, and then calculated HR and FAR. To control for possible co-variation of $l|l$ and test position, we constructed a “local” FAR separately for adjacent and remote lags by including data from new items that were contiguous with adjacent or remote old-item pairs, respectively (Murdock & Kahana, 1993). HR was significantly higher for successive probes from adjacent list positions than for successive probes from remote positions for the three most conservative confidence levels, paired-sample $t(90) = 2.19, 3.00,$ and $2.52, p < .05$; Cohen’s $d_s = 0.46, 0.63,$ and $0.53$, respectively.

We further examined whether level of confidence mediated the recognition advantage for temporally proximate items. Let
We next examined the effects of $r$ on highest-confidence HR for items preceded by lower-confidence “old” responses. We found no effect of $|r|$ on HR for probes preceded by responses 1 through 5, $t(90) = 0.06–1.04$, all $p > .3$. Noting that none of these responses showed a slope different from zero (and that these slopes did not differ from one another), we collapsed prior responses 1 through 5 into a single category (Fig. 3, open circles). The mean slope collapsed across response categories 1 through 5 was not different from zero ($0.000 \pm 0.004$). Our power to detect an effect of the same magnitude as that observed for the highest-confidence responses was .91 (two-tailed). A total of 3,929 data points entered into this analysis of response categories 1 through 5. Note that the slope of the linear relation between $|r|$ and HR for $O_j$ was significantly smaller for these lower-confidence responses to $O_i$ than for the highest-confidence responses to $O_i$, $t(90) = 2.67$, $p < .02$, $d = 0.56$.

We next considered, and rejected, an encoding-based account of the associative effects shown in Figure 3. Suppose that during study, the subject can be in either a good or a bad encoding state. Suppose further that these states have some “inertia,” so that it takes some time to switch from one state to another. This latter assumption is plausible—examples of such long-range correlations have been reported in behavioral time series data (Van Orden, Holden, & Turvey, 2003; Wagenmakers, Farrell, & Ratcliff, 2005). Under these two assumptions, items that receive a highest-confidence response during testing would likely have been encoded in the good-encoding state. Because this good state is hypothesized to persist for some time, items from nearby list positions would also be likely to have been learned in the good-encoding state. Correlated encoding states could therefore lead to a pattern of results like that observed in Figure 3. As we show next, however, the implications of this correlated-encoding hypothesis and the temporal-retrieval hypothesis are readily distinguishable.

The correlated-encoding hypothesis predicts that contiguity effects should be observed even for pairs of items that are not tested successively. We tested this hypothesis by regressing HR on $|r|$ for randomly chosen pairs of old-item test probes that were not tested successively. To ensure that our findings could not be due to a particular randomization, we repeated the analyses on 10,000 randomizations, each time evaluating the mean regression slope (across subjects) as in the original analysis. According to the correlated-encoding hypothesis, we should have seen the same effect of $|r|$ on HR as shown in Figure 3. That is, an item studied in a list position proximate to that of a “sure old” item should have tended to be recognized as “sure old,” even though the two items were not tested successively.

The mean regression slope for these nonsuccessively tested pairs, $-0.004$, was significantly different from zero (the standard deviation across shuffles was .003, $p < .001$). We can therefore conclude that there was small, but significant, autocorrelation in the quality of encoding. However, this effect was not nearly large enough to account for the observed slope for successively tested pairs: $-0.014 \pm 0.004$. Indeed, the slope for successively tested pairs was 2.9 standard deviations away from the slope for nonsuccessively tested pairs ($p < .002$). Thus, the effect shown in Figure 3 depends critically on test order and reflects a retrieval phenomenon.

**GENERAL DISCUSSION**

The present results show that when two old items are tested successively, memory for the second is better if it was initially...
presented in temporal proximity to the first (Fig. 2). This
tendency, however, is wholly attributable to cases in which the
first item receives a highest-confidence response (Fig. 3). These
highest-confidence “old” responses may be considered to reflect
successful recollection of specific attributes of the encoding
episode, whereas lower-confidence “old” responses are assumed
to reflect the familiarity of an item whose attributes are not
recollected (Sherman, Atri, Hasselmo, Stern, & Howard, 2003;
Yonelinas, 1999).

Our finding of temporal associative effects in item recognition
suggests that recollection of an item not only retrieves detailed
information about the item tested, but also retrieves information
about the item’s neighbors. On the basis of episodic recall data,
temporal associations have been taken as evidence for rehearsal
processes orchestrated by a working or short-term memory
system. One might ask whether such rehearsal processes could
explain the associative tendencies in item-recognition demonstrated here. Although our pictures were not likely to have been
coded in a primarily verbal manner, it is nonetheless possible
that verbal recoding was operative for at least a subset of our
stimuli. Thus, our findings may be a consequence of verbal
rehearsal as characterized by the Search of Associative Memory
(SAM) model of Shiffrin and his colleagues (e.g., Gillund &
Shiffrin, 1984) and by models of the phonological loop
(e.g., Burgess & Hitch, 1999). If one could devise stimuli that
would thoroughly resist verbal coding, an account based on
visuospatial working memory could still potentially predict as-
sociative effects.

Theoretically, it would be appealing if the effects observed
here in item recognition could be related to similar associative
tendencies that have been carefully documented in episodic-
recall tasks. In free recall, successively recalled items tend to
have been studied in nearby list positions (Kahana, 1996). By
analogy to the recency effect, in which items near in time to the
end of the list are remembered better than earlier items, we
(Howard & Kahana, 1999) have referred to associative effects in
free recall as illustrating a lag-recency effect, as they reveal a
preference for recalling items presented near in time to the just-
recalled item. Very similar lag-recency effects are observed in
serial recall (Klein, Addis, & Kahana, 2005; Kahana & Caplan,
2002; Raskin & Cook, 1937). A difference between the results
observed here and the lag-recency effects observed in both free
and serial recall is that in both recall tasks, subjects show a
strong bias for making recall transitions in the forward direction.
The fact that our recognition data do not show any reliable
asymmetry points to a potential difference between the associ-
ative processes underlying recognition and recall tasks.

The SAM model offers one possible framework in which to
unify all of these associative effects. In SAM, associations are
formed between nearby items during encoding by virtue of their
co-occurrence in short-term store. In episodic tasks, the just-
recalled item serves as a cue for subsequent recalls, resulting in
associative effects. In recognition testing, it is possible that
associative effects across subsequent test items, such as the
effects observed here, could result if prior test items, stored in
short-term memory, can contribute to a recognition decision for
the current test.

Another possible approach to describing associative effects
across paradigms would be to use a model based on retrieved
temporal context. Context has long been used as an explanatory
concept to describe changes in memory states over time (An-
derson & Bower, 1972; Mensink & Rauijmakers, 1988; Yntema
& Trask, 1963). By assuming that when an item is recalled it
retrieves its encoding context, we (Howard & Kahana, 2002)
were able to explain the lag-recency effect in free recall, and its
approximate invariance across varying interitem delays.
Because retrieved context of a given item overlaps with the
encoding context of nearby items, associative effects can occur.
This approach constitutes a departure from traditional accounts
of association that assume direct item-to-item connections.

Retrieved temporal context has also been proposed as an
explanation for a broad range of effects in item recognition. In
their context-noise recognition model, Dennis and Humphreys
(2001) proposed that a test probe retrieves contextual elements
stored when the item was previously studied. For old test probes,
this retrieved context will include elements from the study list,
and as well as preexperimental exposures to the word; for new test
probes, this retrieved context will include only contextual ele-
ments from preexperimental exposures. For each test probe, the
retrieved contextual information is compared with a context
vector representing the entire list. Dennis and Humphreys
showed that their context-noise model accounts for a number of
benchmark findings in item recognition, including the null
list-strength effect (Ratcliff, Clark, & Shiffrin, 1990; but see
Norman, 2002), the mirror effect (Glazer & Adams, 1990), and
findings from Jacoby’s (1992) process dissociation procedure.
Notably, the context-noise model predicts that context variabil-
ity—the number of preexperimental contexts in which a word
has been experienced—should have an effect on recognition
performance beyond the effect of word frequency. In a study
consistent with this prediction, Steyvers and Malmberg (2003)
found better recognition performance for words appearing in
fewer preexperimental contexts as compared with words appear-
in in a larger number of preexperimental contexts.

A modified version of Dennis and Humphreys’ (2001) context-
noise model could reproduce the associative effects in item
recognition reported here. First, the context retrieved by items
would have to change gradually over time. This would be nec-
esary to obtain the gradually sloping contiguity effect observed
here (Fig. 3). Second, information recovered from one test probe
would have to persist to contribute to the recognition decision on
the subsequent probe. Third, some type of variability in con-
textual retrieval would have to be introduced to account for the
distinction between recollection and familiarity that critically
determines whether associative effects are observed (Fig. 3).
The first two of these modifications have already been proposed
in developing a model of the lag-recency effect in free recall (Howard, 2004; Howard & Kahana, 2002) and would be relatively straightforward revisions to Dennis and Humphreys’ model.

In summary, our view is that the information retrieved by a successful recollection includes the temporal context from the time the item was first presented (Dennis & Humphreys, 2001). Temporal context changes gradually over time (Howard & Kahana, 2002) and persists long enough after retrieval to contribute to the recognition decision on the next probe item. Taken together, the temporally defined associations in recognition revealed here and the ubiquitous observation of temporally defined associations in recall (Howard & Kahana, 1999; Kahana, 1996; Kahana & Caplan, 2002; Raskin & Cook, 1937) suggest that recovery of temporal context is a fundamental process central to episodic memory. Furthermore, our finding that associative effects in recognition are limited to recollective responses adds to a growing body of neuroscientific evidence (reviewed by Rugg & Yonelinas, 2003) suggesting that recollection and familiarity are mediated by different brain mechanisms or structures.

Acknowledgments—We acknowledge support from National Institute of Mental Health (NIMH) Grants MH55687 and MH61975 to the University of Pennsylvania and NIMH Grant MH61492 to Syracuse University. We wish to thank Endel Tulving, Dan Schacter, Mike Humphreys, Ken Norman, and Andy Yonelinas for helpful discussions concerning this work. We also thank Madhura Phadke and Radha Modi for assistance with data collection and Jim Steinhart for photographic materials.

REFERENCES


(Received 7/30/04; Revision accepted 4/18/05; Final materials received 6/1/05)