When does a different environmental context make a difference in recognition? A global activation model

KEVIN MURNANE and MATTHEW P. PHELPS
Pennsylvania State University, University Park, Pennsylvania

The effects on recognition of changes in environmental context between learning and test are examined. A context effect occurs when memory tests that take place in an environmental context that is different from the learning context produce consistent differences in performance. A formal model of context-dependent recognition within a global activation framework is presented. The model generates the predictions that (1) context effects will be present when items are tested in a new context that was not seen during learning and (2) context effects will be absent or very small when items are tested in a context that was experienced during learning but that differs from the context in which the test item was learned. Both predictions were verified in an experiment that varied the nature of the different-context test within subjects. Implications for research concerned with context-dependent recognition are discussed.

Prior studies that have tested recognition memory in an environmental context that is different from the learning context have produced inconsistent results. In a number of studies (Geiselman & Bjork, 1980; Geiselman & Glenny, 1977; Murnane & Phelps, 1993; Smith, 1986; Smith & Vela, 1992), changes in environmental context between learning and test have affected recognition performance; in other studies (Fernandez & Glenberg, 1985; Godden & Baddeley, 1980; Jacoby, 1983; Smith, 1982; Smith, Glenberg, & Bjork, 1978), effects of changes in context have not been observed (see Smith, 1988, for a review). In this paper, we present a formal model of the effects of context on recognition and use the model to point out a critical design feature of context-change experiments that may be responsible for some of the discrepancies in prior results.

Context information can be defined as any information that is present in the processing environment and incidental to the cognitive task being performed. Encoded context information can fulfill an important memory function by identifying different stored events that involve the same memory target as unique episodes. For example, two memories about the same person can be distinguished on the basis of differences in the context information that each contains. Many memory theories incorporate the idea, expressed by Tulving (1983; Tulving & Thompson, 1973) as the encoding specificity prin-

ciple, that a match between the information stored in memory and the information in a cue that is used to access memory promotes successful retrieval. Thus, a cue-to-memory match in the context information that identifies a unique episode is expected to facilitate retrieval of that episode. In accord with these ideas, matched context information in memory and the retrieval cue has been cited as an important factor in the successful retrieval of specific learning episodes (Gillund & Shiffrin, 1984; Humphreys, Bain, & Pike, 1989), and mismatched context information at learning and test has been cited as an important cause of forgetting (Bjork & Richardson-Klavehn, 1989; Mensink & Raaijmakers, 1988, 1989; Raaijmakers & Shiffrin, 1981; Smith, 1988).

With certain notable exceptions (e.g., Fernandez & Glenberg, 1985; see Smith, 1988, for a review), the results obtained when memory is tested with free or cued recall support this view of the importance of matching environmental context information at learning and test. However, as noted above, the results from recognition studies are less clear. Several hypotheses have been advanced to explain the inconsistent recognition results. According to relative cue strength hypotheses like the outshining hypothesis (Smith, 1988, in press; Smith & Vela, 1986; Vela, 1989) the test item in the typical recognition test is such a strong memory cue that it masks or "outshines" any effects that context information might have. According to the reinstatement hypothesis of Bjork and Richardson-Klavehn (1989), subjects routinely mentally reinstate the learning context in different-context test conditions, thereby eliminating any effects of testing in a different context. Although empirical support exists for both hypotheses, neither has yet succeeded in explaining the entire pattern of context-dependent recognition findings.

In the typical experiment designed to examine environmental context-dependent recognition, subjects learn a set of items in one or several learning contexts. Individual items are then tested either in the environmental context in which they were learned or in a different context. An environmental context effect is indicated by consistent differences in performance between same- and different-context test conditions. The most common manipulation of environmental context used in previous research has been to have learning take place in one room and testing take place in the same or a different room (E. Eich, 1985; Fernandez & Glenberg, 1985; Jacoby, 1983; Smith, 1979, 1985, 1986; Smith et al., 1978). Variants of this approach have involved learning lists of words underwater and undergoing memory tests on land (Godden & Baddeley, 1980) or on a boat (Emmerson, 1986). Other studies have manipulated voice (Geiselmann & Bjork, 1980; Geiselmann & Glenny, 1977) or background color (Dulsky, 1935; Weiss & Margolius, 1954). Studies that have used computer presentation of learning and test stimuli have manipulated screen colors and locations either alone (Murnane & Phelps, 1993) or in combination with the shape of a border around the stimulus item and an auditory tone presented with the item (Wright & Shea, 1991).

The task of making theoretical sense out of the tangle of contradictory empirical findings is complicated by the wide variety of dependent measures, types of recognition test, and methods used to manipulate environmental context and to instantiate the different-context test condition used in prior studies. Rather than attempt to examine all possible combinations of design features, we adopt a theoretical stance on recognition memory and examine the implications for experimental design and consequent empirical predictions.

Context Effects and Global Activation Theories of Memory

In general, global activation theories of recognition (SAM, Gillund & Shiffrin, 1984; Minerva 2, Hintzman, 1986, 1988; Matrix model, Humphreys et al., 1989; TODAM, Murdock, 1982, 1989; CHARM, J. M. Eich, 1982, 1985; and many neural net models) propose that recognition involves matching information in a test cue against a large set of items in memory. The test cue activates each item in the memory set, and the total output from the memory, called global activation, serves as the basis for the recognition decision. The magnitude of the output is a function of the match between information in the cue and information in memory; matched information produces higher levels of activation than mismatched information and an increased likelihood that the test item will be recognized as having been previously experienced.

The global matching theories differ in the assumptions that they make about the form of the memory representation (see Hintzman, 1990, and Raaijmakers & Shiffrin, 1992, for reviews). In SAM (Gillund & Shiffrin, 1984) and Minerva 2 (Hintzman, 1986, 1988), individual items are encoded in separate memory representations. At test, the recognition cue activates each item separately and the individual item activations are summed to give the global activation that serves as the basis for the recognition decision. In the Matrix model (Humphreys et al., 1989), TODAM (Murdock, 1982, 1989), and CHARM (J. M. Eich, 1982, 1985), the memory representation is composite. Individual items are added to the composite memory at encoding. At test, the recognition cue activates the composite memory to produce a global activation. In the composite models, the magnitude of the global activation can be calculated as a sum of activations of individual items, even though, strictly speaking, items are not individually represented in the composite memory. For ease of exposition, we will discuss all the global activation theories in terms of the activation of individual memory representations. The critical point for our analysis is that all the theories assume that context information and information about the to-be-remembered item (hereafter, item information) are contained both in memory and in the test cue and that activation in response to both types of information contributes to the global activation on which recognition decisions are based.

The theories also differ in the way activation is calculated from the (mis)match between the item and context information in the test cue and in memory. Let M represent global activation in response to a test item. In all the current global matching theories, global activation is calculated as the sum of the activations of individual items—that is,

$$M = \sum_{j=1}^{K} f(I_j, C), \qquad (1)$$

where I and C are positive real numbers greater than zero, I represents the strength of the match between item information in the cue and memory, C represents the strength of the match between context information in the cue and memory, K is the number of items in memory, and f is a strictly increasing function of both I and C. For purposes of illustration, suppose that there is only one item in memory. In this case, global activation is the activation of a single item represented as

$$M = f(I, C). \qquad (2)$$

The global activation theories differ in the nature of the activation function, f. In many of the composite storage theories, global activation is calculated as the dot product of the feature vectors (TODAM and CHARM) or matrices (Matrix) that represent the test cue and memory. In the separate storage theories, global activation is a nonlinear function of both I and C. In SAM, the activation function is negatively accelerated. Activation of a single item in memory is given by

$$M = I^{w_i} \times C^{w_c}, \qquad (3)$$

where w_i and w_c are weights assigned to the item and context cues, respectively. The weights are constrained

to sum to 1.0. In Minerva 2, the activation function is positively accelerated. Items are represented as feature vectors, and activation of a single item in memory is given by

$$M = S(I,C)^3, \qquad (4)$$

where $S(I,C)$ is a similarity function defined over the (mis)matching item and context elements of the vectors that represent the test cue and memory. The nonlinear models make different predictions from certain versions of the dot product models under circumstances that will be discussed below.

Although our immediate goal is to examine context-dependent recognition in the set of global activation theories mentioned above, the conclusions we reach can be generalized to all global activation theories in which the activation of a single item in memory is a strictly increasing function of both the item and context match strengths. Let a subscripted H represent a high level of activation produced by matched information in memory and the test cue, and let a subscripted L represent a low level of activation produced by mismatched information. There are four activation quantities shown in Table 1 that contribute to global activation. On the assumptions that matched information in the test cue and memory results in higher levels of activation than mismatched information, and that the activation function is strictly increasing with both I and C, the following ordinal ranking of the four activation quantities is straightforward:

$$f(I_{\rm H},C_{\rm H}) > f(I_{\rm H},C_{\rm L}) \; ? \; f(I_{\rm L},C_{\rm H}) > f(I_{\rm L},C_{\rm L}).$$

The relationship between $f(I_{\rm H},C_{\rm L})$ and $f(I_{\rm L},C_{\rm H})$ is unknown because it depends on assumptions made about relative levels of activation in response to context and item information, respectively, and about the relative strength of both types of information in memory.

Consider a list of K items divided among J learning contexts with N items learned in each context (thus, $K = JN$). At test, items are tested either in one of the learning contexts or in a new context that was not experienced during learning. There are four types of test item: targets tested in the context in which they were learned, targets tested in a new context, distractors tested in one of the learning contexts, and distractors tested in the new context. We call a general experimental design of this type an AB–X paradigm; items are learned in either one or several contexts A (and B ...), and different-context tests take place in an unstudied context, X.

Table 1
Activation Quantities for Context-Change Experiments

Item	Context	
	Match	Mismatch
Match	$f(I_{\rm H},C_{\rm H})$	$f(I_{\rm H},C_{\rm L})$
Mismatch	$f(I_{\rm L},C_{\rm H})$	$f(I_{\rm L},C_{\rm L})$

Note—$I_{\rm H}$, high level of item activation; $I_{\rm L}$, low level of item activation; $C_{\rm H}$, high level of context activation; $C_{\rm L}$, low level of context activation; f is the activation function.

Assume that, in a typical list-learning experiment, only items that have been encoded from the study list contribute to the global activation that is the basis for recognition. Global activation in response to same- (subscripted s) and different- (subscripted d) context targets (subscripted T) and distractors (subscripted D) in an AB–X paradigm can then be represented as follows:

$$M_{\rm Ts} = f(I_{\rm H},C_{\rm H}) + (N-1)f(I_{\rm L},C_{\rm H}) + (K-N)f(I_{\rm L},C_{\rm L})$$
$$(5)$$

$$M_{\rm Td} = f(I_{\rm H},C_{\rm L}) + (K-1)f(I_{\rm L},C_{\rm L}) \qquad (6)$$

$$M_{\rm Ds} = Nf(I_{\rm L},C_{\rm H}) + (K-N)f(I_{\rm L},C_{\rm L}) \qquad (7)$$

$$M_{\rm Dd} = Kf(I_{\rm L},C_{\rm L}). \qquad (8)$$

To see how different activation quantities contribute to global activation, consider the case for targets tested in their learning context (Equation 5). The item and context information in the test cue match the item and context information in the memory representation of the target producing the single $f(I_{\rm H},C_{\rm H})$ match strength. The remaining $(N-1)$ items that were learned in the same context as the target match the context information and mismatch the item information in the test cue producing the $(N-1)f(I_{\rm L},C_{\rm H})$ match strengths. Everything else in memory mismatches both the item and context information in the test cue producing the $(K-N)f(I_{\rm L},C_{\rm L})$ match strengths. Equations 6, 7, and 8 can be analyzed in a similar manner.

Comparisons of global activations in same- and different-context test conditions is facilitated by moving the $(K-N)f(I_{\rm L},C_{\rm L})$ factors that are common to all four equations to the left side of the equations. This leaves the components of the global match that constitute the differences in activation between the four types of probe (symbolized with a superscripted *):

$$M_{\rm Ts}{}^* = f(I_{\rm H},C_{\rm H}) + (N-1)f(I_{\rm L},C_{\rm H}) \qquad (5a)$$

$$M_{\rm Td}{}^* = f(I_{\rm H},C_{\rm L}) + (N-1)f(I_{\rm L},C_{\rm L}) \qquad (6a)$$

$$M_{\rm Ds}{}^* = Nf(I_{\rm L},C_{\rm H}) \qquad (7a)$$

$$M_{\rm Dd}{}^* = Nf(I_{\rm L},C_{\rm L}). \qquad (8a)$$

Examination of Equations 5a–8a makes it readily apparent that global activation is predicted to be higher in same-context test conditions than in different-context test conditions in an AB–X paradigm. For targets (Equations 5a and 6a) the single $f(I_{\rm H},C_{\rm H})$ strength and the $(N-1)f(I_{\rm L},C_{\rm H})$ strengths for same-context tests are, respectively, greater than the $f(I_{\rm H},C_{\rm L})$ and the $(N-1)$ $f(I_{\rm L},C_{\rm L})$ strengths for different-context tests. Likewise, for distractors (Equations 7a and 8a), the $Nf(I_{\rm L},C_{\rm H})$ strengths for same-context tests are greater than the $Nf(I_{\rm L},C_{\rm L})$ strengths for different-context tests.

One consequence of the global activation approach to recognition is that context information, taken alone, is predicted to make the same contribution to global activation for studied and unstudied items. When tests occur in one of the learning contexts, there are N cases of con-

text match and $K-N$ cases of context mismatch (see Equations 5 and 7), and when tests occur in a new context, there are K cases of context mismatch (see Equations 6 and 8). The fact that context contributes equally to activation for targets and for distractors has two important methodological implications. First, test context should be manipulated within lists. If test context is manipulated between lists, mean activations for both targets and distractors will be greater on the same-context list than on the different-context list, with the result that differences in performance in same- and different-context test conditions may be masked by shifts in response criteria between lists. Second, care must be taken to use an appropriate dependent measure. Standard measures of recognition performance such as d' are based on the difference between mean activation in response to targets and distractors. Because (mis)matched context makes the same contribution to these two means, changes in test context may not affect the standard measures. The individual global activation theories are capable of predicting that context change will have a negative effect, no effect, or a positive effect on d', depending on the activation function used in each theory and on assumptions made about the relationship between the mean and the variance of global activation (for a fuller discussion of these issues, see Murnane & Phelps, 1993).

A fundamental tenet of the global activation framework for recognition is that increases in global activation result in an increased likelihood of identifying the test item as "old." Thus, clear predictions are made about the hit rate (HR) and false alarm rate (FAR); both are predicted to be higher in same-context test conditions than in different-context test conditions. To simplify the discussion, assume that activation in response to both target and distractor tests is normally distributed. The HR and FAR predictions then depend on the assumptions that (1) variances are equal for same- and different-context tests, and (2) subjects do not systematically shift response criteria between same- and different-context tests. If subjects are systematically more conservative in setting response criteria on the different-context tests in the AB–X paradigm, one would expect the same pattern of results as that predicted by global matching theories. In the experiment reported below, several steps are taken to minimize the possibility that changes in performance with changes in test context are caused by subject-initiated shifts in response criteria. First, test context is manipulated within subjects and within lists, and same- and different-context tests are randomly intermixed during testing. Thus, criterion shifts would have to occur on a trial-by-trial basis. Second, test context is not predictive of correct response. Third, performance feedback is not given. Thus, subjects are given no external motivation to systematically shift criteria in response to changing test conditions. Manipulation of test context within lists is also assumed to produce equal variances for same- and different-context tests.

Comparison of the AB–X and AB–A Paradigms

Many published studies have used an AB–X paradigm (e.g., Godden & Baddeley, 1980; Murnane & Phelps, 1993; Smith, 1982, 1985; Smith et al., 1978). Using a within-list manipulation of test context and the appropriate dependent measure, Murnane and Phelps (1993) showed that the context effects predicted by global activation theories were present in a series of three studies that used associative recognition and two studies that used item recognition.

Several studies (e.g., Humphreys, Pike, Bain, & Tehan, 1988; Smith, 1985) have employed an experimental design we call an AB–A paradigm. In an AB–A paradigm, learning takes place in at least two contexts, A and B, and testing takes place in one of the learning contexts, say A. Items originally learned in A are thus tested in the same context, and items originally learned in B are tested in a different context.

How do predictions change if context-dependent recognition is examined using an AB–A paradigm? Comparative predictions cannot be made about FAR, because all distractors are tested in one of the learning contexts; that is, all distractor activations are predicted by Equation 7.

Let M_{TAA} represent global activation for targets tested in the learning context and M_{TAB} represent global activation for targets tested in a different context. Then, global activations for targets in an AB–A paradigm are given by Equations 9 and 10.

$$M_{TAA} = f(I_H, C_H) + (N-1)f(I_L, C_H) + (K-N)f(I_L, C_L)$$
(9)

$$M_{TAB} = f(I_H, C_L) + (N)f(I_L, C_H) + (K-N-1)f(I_L, C_L).$$
(10)

Note that Equation 9 is the same as Equation 5; that is, global activation for same-context targets is identical in the AB–A and AB–X paradigms. The difference between the two paradigms is caused by differences in global activation in response to different-context tests. Moving the $(K-N-1)f(I_L, C_L)$ factors and the $(N-1)$ $f(I_L, C_H)$ factors that are common to Equations 9 and 10 to the left side of the equations leaves the components of the global match that constitute the difference in activation between same- and different-context tests:

$$M_{TAA}^* = f(I_H, C_H) + f(I_L, C_L)$$
(9a)

$$M_{TAB}^* = f(I_H, C_L) + f(I_L, C_H).$$
(10a)

Comparison of Equations 9a and 10a with Equations 5a and 6a shows two critical differences in the global activation for targets in the AB–A and AB–X paradigms. First, in the AB–X paradigm, the increased activation caused by matched context information, C_H, increases global activation for same-context tests under all circumstances (Equations 5a and 6a). However, in the

AB–A paradigm, matched context information increases global activation for same-context tests when item information matches (Equation 9a) and increases global activation for different-context tests when item information mismatches (Equation 10a). Second, if N is greater than 2, more individual factors contribute to the magnitude of the global sum in the AB–X paradigm.

How do these differences affect predictions of environmental context effects in the AB–A paradigm? The differences in global activation between same- and different-context target tests in the AB–X and AB–A paradigms are given by Equations 11 and 12, respectively:

$$\Delta M_{\text{TAB-X}} = f(I_H, C_H) - f(I_H, C_L) +$$
$$(N-1)[f(I_L, C_H) - f(I_L, C_L)] \quad (11)$$

$$\Delta M_{\text{TAB-A}} = f(I_H, C_H) - f(I_H, C_L) +$$
$$f(I_L, C_L) - f(I_L, C_H). \quad (12)$$

Let $f(I_H, C_H) - f(I_H, C_L)$ equal x, which will be a positive number because $f(I_H, C_H)$ (both item and context information match) is greater than $f(I_H, C_L)$ (item information matches and context information mismatches); let $f(I_L, C_H) - f(I_L, C_L)$ equal y, which will be a positive number because $f(I_L, C_H)$ is greater than $f(I_L, C_L)$; and let $f(I_L, C_L) - f(I_L, C_H)$ equal z, which will be a negative number because $f(I_L, C_L)$ is less than $f(I_L, C_H)$. Then, Equations 11 and 12 can be rewritten as

$$\Delta M_{\text{TAB-X}} = x + (N-1)y \quad (11a)$$

$$\Delta M_{\text{TAB-A}} = x - z. \quad (12a)$$

Clearly, context effects are predicted to be smaller in an AB–A paradigm than in an AB–X paradigm.

An interesting special case exists in which context effects are predicted to be absent in an AB–A paradigm. Reconsider Equations 9a and 10a, which give the components of the global match that constitute the difference in global activation between same- and different-context tests in an AB–A paradigm:

$$M_{\text{TAA}}^* = f(I_H, C_H) + f(I_L, C_L) \quad (9a)$$

$$M_{\text{TAB}}^* = f(I_H, C_L) + f(I_L, C_H). \quad (10a)$$

If the activation function is simple addition, Equations 9a and 10a are equivalent. Removing parentheses and rearranging terms in Equation 10a gives

$$M_{\text{TAA}}^* = M_{\text{TAB}}^* = I_H + C_H + I_L + C_L. \quad (13)$$

Thus same- and different-context targets produce the same magnitude of global activation, and context effects are predicted to be absent. In general, let a, b, c, and d represent different match strengths. Context effects are predicted to be absent for any case in which

$$f(a, b) + f(c, d) = f(a, d) + f(c, b). \quad (14)$$

We refer to functions for which Equation 14 is true as functions that meet the match rearrangement criterion. This is an interesting case, because identifiable variants

of models that use a dot product activation function meet the criterion. For example, global activation in TODAM is given by the dot product between a vector representing memory and a vector representing the test cue. If context and item information in the test cue are represented as separate subcomponents of the cue vector, TODAM meets the match rearrangement criterion and predicts that context effects will be absent in an AB–A paradigm. Note that it is not necessary that context and item information be represented as separate subcomponents of the vector representing memory. If, however, context and item information are not kept separate in the test cue, as is normally the case in CHARM and the Matrix model, then the match rearrangement criterion is not met, and context effects are predicted to be smaller in an AB–A paradigm than in an AB–X paradigm.

To summarize the foregoing qualitative predictions, context effects for both targets and distractors are predicted when recognition memory is tested using an AB–X paradigm. When an AB–A paradigm is used, context effects for targets are predicted to be either absent or smaller in magnitude than the predicted effect under equivalent conditions in an AB–X paradigm. Predictions comparing context effects for distractors in the two paradigms cannot be made because all distractors are tested in a learning context in an AB–A paradigm. These predictions were tested in an experiment that manipulated test context (same vs. different) and test paradigm (AB–A vs. AB–X) completely within subjects. The general experimental paradigm was a variant of the paradigm used by Murnane and Phelps (1993).

METHOD

Subjects

Subjects were 60 undergraduate volunteers enrolled in introductory psychology courses who earned course credit in return for their participation.

Materials

Word pairs and distractors were generated from a set of 720 high-frequency nouns (Francis & Kučera, 1982). Paired words and distractors were randomly chosen and randomly assigned to conditions for each subject.

Design

In order to minimize the likelihood that subjects would associate items studied in one learning context with the other learning context, study items consisted of pairs of words, and subjects were instructed to associate the members of each pair during learning. Ten lists of 24 study pairs followed by 48 single-item recognition tests were presented during the experimental session. Targets were 24 single words chosen randomly from either the first or the second member of each study pair. Distractors were 24 new words. A 30-sec arithmetic task was presented between the study and test lists to prevent retrieval from short-term or working memory (Peterson & Peterson, 1959).

Study pairs were equally divided between two learning contexts. Context was operationally defined as a unique combination of foreground color, background color, and location on a computer screen. The two learning contexts had the following attributes. In Context A, light green words were presented on a magenta back-

ground in the upper left corner of the screen. In Context B, red words were displayed on a gray background in the lower right corner of the screen. In Context X, the different test context in the AB–X paradigm, black words were presented on a cyan background in the center of the screen. These contexts were very similar to three of the contexts used by Murnane and Phelps (1993).

Test context was manipulated within subjects and within lists. Words were tested either in the same context in which they were learned or in a different context. Test paradigm was manipulated within subjects and between lists. Five lists from each paradigm were presented. On every list, six targets that were learned in Context A and six targets that were learned in Context B were tested in the same context in which they were learned. On AB–X lists, the remaining six targets from each learning context were tested in Context X. Half of the distractors were equally divided between Contexts A and B, and half were tested in Context X. On AB–A lists, six targets from each learning context were tested in the opposite learning context (e.g., a word learned in A would be tested in B). Distractors were equally divided between Contexts A and B; thus, there were no different-context tests for distractors in the AB–A paradigm.

Procedure

Subjects were seated in individual booths with personal computers, were told the nature of the memory task, and were instructed to memorize the word pairs by forming associations between the words in each pair. They were told to respond "old" or "new" to each test word on the basis of whether or not the word had appeared in the study list, regardless of the screen location or color. A short practice list was given to familiarize subjects with the procedure.

Subjects initiated the presentation of each list by pressing the space bar. Study pairs were presented for 3 sec each, followed by 300 msec of blank screen. During the arithmetic task, seven single-digit numbers were presented one at a time for 3 sec each. A plus sign was displayed for 1 sec between digit presentations. After presentation of the last digit, an equals sign appeared, and subjects were allowed 7 sec to enter the total. Test words were presented until a response was given or for a maximum of 3 sec, followed by 300 msec of blank screen. Subjects entered their recognition decisions by pressing either the "D" or the "K" key on the computer keyboard. Mapping of recognition responses to keys was counterbalanced across subjects. Presentation order of lists, study pairs, and test words was randomized for each subject.

RESULTS

Analyses were conducted with mean HRs and FARs as dependent measures. Context effects were measured as the difference between means in same-context and in different-context test conditions. A difference score greater than 0.0 indicates a context effect. Planned comparisons using within-subject t tests were conducted on the difference-between-means scores to examine the context effects for each test paradigm. The means and the results of the planned comparisons are given in Table 2.

As predicted, a significant context effect was found for both targets and distractors in the AB–X paradigm. These results replicate the context effects reported by Murnane and Phelps (1993). Also as predicted, a small context effect that was not significant was found for targets in the AB–A paradigm.

Table 2
Mean Hit and False Alarm Rates and Planned Comparisons

Test Paradigm	HR_S	HR_D	$HR_S - HR_D$	S_D	t	df	p
AB–X	.742	.696	.046	.010	4.40	59	<.0005
AB–A	.735	.718	.017	.010	1.70	59	.095
	FAR_S	FAR_D	$FAR_S - FAR_D$				
AB–X	.267	.234	.033	.012	2.87	59	.006
AB–A	.249	N/A					

Note—HR_S, hit rate for same-context tests; HR_D, hit rate for different-context tests; $HR_S - HR_D$, the mean difference; FAR_S, false alarm rate for same-context tests; FAR_D, false alarm rate for different-context tests; $FAR_S - FAR_D$, the mean difference; S_D, standard error of the mean difference; p, two-tailed probability value.

A repeated-measures ANOVA was conducted with mean HRs as the dependent measure. There was a significant main effect of test context $[F(1,59) = 17.97, p < .0005]$. The test context $\times$ test paradigm interaction was marginally nonsignificant $[F(1,59) = 3.89, p = .053]$. The main effect of test paradigm was not significant $(p = .3)$.

Friedman (1968) recommended the use of point-biserial correlations between independent and dependent variables as a measure of effect size. This procedure was used to compare the magnitude of the observed context effect $(HR_S - HR_D)$ in the two test paradigms. Test context accounted for 24.7% of the variance $(r = .50)$ in the AB–X test paradigm, but only 4.7% of the variance $(r = .22)$ in the AB–A paradigm.

DISCUSSION

The results of this experiment confirmed the predictions made by global matching theories that context effects would be present when recognition performance is measured using an AB–X paradigm and absent or very small when performance is measured using an AB–A paradigm. Taken in conjunction with the results reported in Murnane and Phelps (1993), these data further clarify the issue of when environmental context effects are to be expected in recognition and provide evidence of the value of the theoretical approach represented in Equations 5–10.

In prior studies, using an AB–A paradigm, Smith (1985) found a small context effect and Humphreys et al. (1988) found no effect. It is difficult to directly compare these results with the predictions generated from Equations 5–10 because both studies tested recognition by asking subjects to pick targets from lists of mixed targets and distractors. This type of recognition test may encourage the use of a cued recall strategy. However, context effects are generally found when memory is tested using cued recall (Smith, 1988). Thus the small magnitude of the observed effects in the Smith (1985) studies and the absence of a context effect in the Humphreys et al. (1988) study are in accordance with expectations based on global matching theories.

The results of the present study are ambiguous with regard to the question of whether or not the activation function meets the match rearrangement criterion. In an AB–A paradigm, models with activation functions that meet the match rearrangement criterion predict an absent context effect and models with activation functions that do not meet the match rearrangement criterion predict a positive effect that is smaller in magnitude than the effect observed in an AB–X paradigm. A small, positive effect that was not statistically reliable ($p = .095$) was found. These results do not provide strong evidence in favor of either type of activation function.

The methodological implications of this study are clear. Failure to find statistically reliable environmental context effects when recognition testing takes place in one of the learning contexts is likely to be an artifact of the experimental paradigm. Care should be taken to ensure that recognition testing takes place in an environmental context that was not experienced during learning if one wishes to obtain reliable context effects in the laboratory.

REFERENCES

BJORK, R. A., RICHARDSON-KLAVEHN, A. (1989). On the puzzling relationship between environmental context and human memory. In C. Izawa (Ed.), Current issues in cognitive processes: The Tulane Floweree Symposium on Cognition (pp. 313-344). Hillsdale, NJ: Erlbaum.

DULSKY, S. G. (1935). The effect of a change in background on recall and relearning. Journal of Experimental Psychology, 18, 725-740.

EICH, E. (1985). Context, memory, and integrated item/context imagery. Journal of Experimental Psychology: Learning, Memory, & Cognition, 11, 764-770.

EICH, J. M. (1982). A composite holographic associative recall model. Psychological Review, 89, 627-661.

EICH, J. M. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. Psychological Review, 92, 1-38.

EMMERSON, P. G. (1986). Effects of environmental context on recognition memory in an unusual environment. Perceptual & Motor Skills, 63, 1047-1050.

FERNANDEZ, A., & GLENBERG, A. M. (1985). Changing environmental context does not reliably affect memory. Memory & Cognition, 13, 333-345.

FRANCIS, W. N., & KUČERA, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin.

FRIEDMAN, H. (1968). Magnitude of experimental effect and a table for its rapid estimation. Psychological Bulletin, 70, 245-251.

GEISELMAN, R. E., & BJORK, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. Cognitive Psychology, 12, 188-205.

GEISELMAN, R. E., & GLENNY, J. (1977). Effects of imagining speakers' voices on the retention of words presented visually. Memory & Cognition, 5, 499-504.

GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-67.

GODDEN, D. R., & BADDELEY, A. D. (1980). When does context influence recognition memory? British Journal of Psychology, 71, 99-104.

HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. Psychological Review, 93, 411-428.

HINTZMAN, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. Psychological Review, 95, 528-551.

HINTZMAN, D. L. (1990). Human learning and memory: Connections and dissociations. Annual Review of Psychology, 41, 109-139.

HUMPHREYS, M. S., BAIN, J. D., & PIKE, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. Psychological Review, 96, 208-233.

HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (1988). Using multilist designs to test for contextual reinstatement effects in recognition. Bulletin of the Psychonomic Society, 26, 200-202.

JACOBY, L. L. (1983). Perceptual enhancement: Persistent effects of experience. Journal of Experimental Psychology: Human Learning, Memory, & Cognition, 9, 21-38.

MENSINK, G.-J. [M.], & RAAIJMAKERS, J. G. W. (1988). A model for interference and forgetting. Psychological Review, 95, 434-455.

MENSINK, G.-J. M., & RAAIJMAKERS, J. G. W. (1989). A model for contextual fluctuation. Journal of Mathematical Psychology, 33, 172-186.

MURDOCK, B. B. (1982). A theory for the storage and retrieval of item and associative information. Psychological Review, 89, 609-626.

MURDOCK, B. B. (1989). Learning in a distributed memory model. In C. Izawa (Ed.), Current issues in cognitive processes: The Tulane Floweree Symposium on Cognition (pp. 69-106). Hillsdale, NJ: Erlbaum.

MURNANE, K., & PHELPS, M. P. (1993). A global activation approach to the effect of changes in environmental context on recognition. Journal of Experimental Psychology: Learning, Memory, & Cognition, 19, 882-894.

PETERSON, L. R., & PETERSON, M. J. (1959). Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.

RAAIJMAKERS, J. G. [W.], & SHIFFRIN, R. M. (1981). Search of associative memory. Psychological Review, 88, 93-134.

RAAIJMAKERS, J. G. [W.], & SHIFFRIN, R. M. (1992). Models for recall and recognition. Annual Review of Psychology, 43, 205-234.

SMITH, S. M. (1979). Remembering in and out of context. Journal of Experimental Psychology: Human Learning & Memory, 5, 460-471.

SMITH, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. Memory & Cognition, 10, 405-412.

SMITH, S. M. (1985). Environmental context and recognition memory reconsidered. Bulletin of the Psychonomic Society, 23, 173-176.

SMITH, S. M. (1986). Environmental context-dependent recognition memory using a short-term memory task for input. Memory & Cognition, 14, 347-354.

SMITH, S. M. (1988). Environmental context-dependent memory. In G. M. Davies & D. M. Thompson (Eds.), Memory in context: Context in memory (pp. 13-34). New York: Wiley.

SMITH, S. M. (in press). Theoretical principles of context-dependent memory. In P. Morris & M. Gruneberg (Eds.), Aspects of memory: Theoretical aspects (2nd ed.).

SMITH, S. M., GLENBERG, A. [M.], & BJORK, R. A. (1978). Environmental context and human memory. Memory & Cognition, 6, 342-353.

SMITH, S. M., & VELA, E. (1986, November). Outshining: The relative effectiveness of cues. Paper presented at the meeting of the Psychonomic Society, New Orleans.

SMITH, S. M., & VELA, E. (1992). Environmental context-dependent eyewitness recognition. Applied Cognitive Psychology, 6, 125-139.

TULVING, E. (1983). Elements of episodic memory. New York: Oxford University Press.

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

VELA, E. (1989, November). Environmental context dependent memory: A meta-analytic review. Paper presented at the meeting of the Psychonomic Society, Atlanta.

WEISS, W., & MARGOLIUS, G. (1954). The effect of context stimuli on learning and retention. Journal of Experimental Psychology, 48, 318-322.

WRIGHT, D. L., & SHEA, C. H. (1991). Contextual dependencies in motor skills. Memory & Cognition, 19, 361-370.