Psychonomic Bulletin & Review
2001, 8 (2), 203-220

Environmental context-dependent memory: A review and meta-analysis

STEVEN M. SMITH
Texas A&M University, College Station, Texas

and

EDWARD VELA
California State University, Chico, California

To address questions about human memory's dependence on the coincidental environmental contexts in which events occur, we review studies of incidental environmental context-dependent memory in humans and report a meta-analysis. Our theoretical approach to the issue stems from Glenberg's (1997) contention that introspective thought (e.g., remembering, conceptualizing) requires cognitive resources normally used to represent the immediate environment. We propose that if tasks encourage processing of noncontextual information (i.e., introspective thought) at input and/or at test, then both learning and memory will be less dependent on the ambient environmental contexts in which those activities occur. The meta-analysis showed that across all studies, environmental context effects were reliable, and furthermore, that the use of noncontextual cues during learning (overshadowing) and at test (outshining), as well as mental reinstatement of appropriate context cues at test, all reduce the effect of environmental manipulations. We conclude that environmental context-dependent memory effects are less likely to occur under conditions in which the immediate environment is likely to be suppressed.

Although life is continually played out within physical contexts, the degree of one's experience of involvement with such environments varies considerably. The attention of a driver engrossed in thought or conversation, for example, can be largely withdrawn from road conditions; in such circumstances, the environment is suppressed (Glenberg, 1997), at least to some degree, while the driver's attention is devoted largely to verbal communication or introspective thought. A sudden road emergency, however, can trigger a heightened degree of attention directed toward the driver's immediate surroundings. Likewise, memories of experiences may vary in how much they are affected by environmental surroundings, both when events are originally experienced and when events are remembered. In some cases, learning and remembering appear to be greatly affected by background environments, and in other circumstances, incidental surroundings influence learning and remembering much less. Under what circumstances are these effects large, and when are they small or nonexistent? Furthermore, do the results of studies of environmental context-dependent memory occur in predictable patterns? In the present review, we offer hypotheses about what those patterns might be, and we demonstrate that our review and meta-analysis are generally consistent with certain of those hypotheses.

In the present study, we examine memory's dependence on the incidental environmental settings in which events are experienced and remembered, focusing on what circumstances affect the likelihood of observing environmental context-dependent memory effects. We review and analyze published results of experimental research on incidental environmental context-dependent memory in humans. We show that this corpus of findings can be explained largely in terms of the principle that the environmental context in which memory testing occurs affects memory in a context-dependent manner, unless task demands encourage subjects to suppress their test environments.

A Brief History of Environmental Context Effects

Experimental studies of the effects of environmental context on various aspects of memory date back at least to Carr (1925), who examined influences of incidental environmental manipulations on maze running in rats. Since that time, numerous studies of contextual effects on human memory have been published, many of which have supported the idea that manipulations of incidental environmental factors affect memory performance. Such studies have been reviewed by Smith (1988, 1994).

Much of the early research on incidental environmental context-dependent memory in humans was done with interference reduction paradigms, in which interfering lists were learned in the same environment or in different ones (e.g., Bilodeau & Schlosberg, 1951; Dallett & Wilcox, 1968; Greenspoon & Ranyard, 1957). It was reported that interference was reduced substantially if the interfering and target lists were learned in different environments. Some doubt was cast on the cause of interfer-

Correspondence should be sent to S. M. Smith, Department of Psychology, Texas A&M University, College Station, TX 77843-4235 (e-mail: sms@psyc.tamu.edu).

ence reduction, however, when Strand (1970) reported that physical disruption between lists caused as much interference reduction as did environmental change. More recent research has obtained interference reduction effects while controlling for physical disruption (e.g., Eckert, Kanak, & Stevens, 1984).

Reinstatement effects, which are now investigated more commonly than interference reduction, refer typically to the finding of better memory when the learning environment is reinstated at test than when testing occurs in a different environment. Under the *different context* condition, in which learning and test occur in different environments, some care is usually taken to use environments that are perceptually distinct from each other. Environmental manipulations have most often been accomplished with room changes, although a variety of other techniques have also been used. Godden and Baddeley (1975, 1980) manipulated context by having scuba divers learn and recall word lists on land or underwater, for example, and Smith and Sinha (1987) used a flotation tank versus a lounge.

Although there were a few early failures to find environmental context effects, these went largely unnoticed (e.g., Farnsworth, 1934; Reed, 1931), and real concern about the reliability of laboratory context effects began to appear in the published literature only more recently. Smith, Glenberg, and Bjork (1978) found that environmental context manipulations did not affect recognition memory, even though there were reliable effects on recall. This noneffect on recognition tests has also been found by others (e.g., Godden & Baddeley, 1980; Jacoby, 1983). Rising concern over the reliability of environmental context effects followed the publication of a set of negative findings by Fernandez and Glenberg (1985), a study by Eich (1985), who found no effect of incidental context on recall, and a study by Saufley, Otaka, and Bavaresco (1985), whose classroom field experiments failed to support the idea of context-dependent memory. These nonfindings underscore the still unresolved empirical, theoretical, and conceptual problems that plague the literature on environmental context-dependent memory. The present review and meta-analysis is in part a response to these published challenges of incidental environmental context-dependent memory effects. Are the effects of incidental environmental context manipulations reliable? If so, what factors appear to modulate the effects?

Rationale for Studying Incidental Environmental Context

Understanding incidental context effects is important for theoretical, empirical, and applied reasons. Many current theoretical models of memory use mechanisms that incorporate contextual associations and the principle of cue-dependent memory (e.g., SAM, Raaijmakers & Shiffrin, 1981; MINERVA 2, Hintzman, 1988; CHARM, J. M. Eich, 1982). The efficacy of such models depends in part on their ability to explain and predict incidental context effects. It is also important to clarify how to manipulate contextual variables in controlled laboratory conditions effectively. A reliable empirical basis for context effects is necessary for one to be able to address theoretical issues concerning memory storage and retrieval. Furthermore, there are applied questions about memory in such domains as aging, eyewitness testimony, psychotherapy, and education. Improving human memory with contextual cues, as well as knowing when contextual manipulations are ineffective, could be potentially beneficial in all of these applied domains.

The central theoretical questions are whether incidental background stimuli are stored in memory, and whether such stimuli can cue memory of materials that are studied contiguously with incidental contexts. Although these theoretical questions are relevant to computational models that rely on contextual associations (e.g., Chappell & Humphreys, 1994; Murnane & Phelps, 1995; Raaijmakers & Shiffrin, 1981), such models generally postulate only that some material serves as context, remaining silent as to whether incidental environmental material is encoded and can cue memory traces. More directly relevant is Glenberg's (1997) theory of environmental suppression, which states that the environment is processed and represented unless efforts to suppress the environment are made to permit conceptual processing. We will return to this theory presently.

Another theoretical issue directly relevant to the issue of context-dependent memory is that of source monitoring (e.g., Johnson, Hashtroudi, & Lindsay, 1993). *Source monitoring* refers to determining the origin of one's own remembered material, and it can involve various types of decision processes related to attributions about memories. An important component of source monitoring is the retrieval of episodic contexts associated with, or "bound to," remembered events (see, e.g., Chalfonte & Johnson, 1996). These episodic contexts are retrieved to help the rememberer respecify the circumstances in which remembered events have occurred, a process that helps the rememberer distinguish among memories that arise from different sources or origins. In the present paper, we are more concerned with the cuing properties of context, rather than the retrieval of contextual information, an important concern in source monitoring. Nevertheless, as we will show, the retrieval of environmental contextual information can benefit the retrieval of episodically associated events. Although source monitoring involves more than the respecification of episodic contexts, it is nonetheless affected by the encoding and retrieval of contextual information, which are important concerns in the present review.

Practical concerns are also relevant to our research question. For example, whether or not classroom exam scores suffer as a function of the testing environment (i.e., the same as the regular classroom vs. an unfamiliar room) is an educational concern. If environmental changes lead to poorer test scores, it is not only of concern to the student, but it should also worry educators, who would likely prefer classroom learning to be independent of learning or training environments. Another practical consideration is eyewitness memory, which may or may not be enhanced by returning the eyewitness to the scene of the witnessed events. Memory in clinical therapy is another application area that could be affected by incidental environmental context-dependent memory. Whether or not reluctant or hidden memories can be elicited by environmental cues is an important concern in clinical psychology. Another emerging practical concern is that of memory deficits associated with aging. One method for alleviating age-related memory deficits could be to use environmental support—that is, to use specially designed environments that provide the elderly with physical cues that serve as memory aids. The importance of these practical concerns highlights the need for a clearer understanding of the effects of incidental environmental context on memory.

There is also the matter of distinguishing our question from others that are already thoroughly researched and are no longer in question. The question that we have posed in the present paper addresses only the effects of global incidental environmental contexts on humans. Related questions that have been studied have included the effects of incidental contexts on nonhuman subjects (see, e.g., Balsam & Tomie, 1985; Carr, 1925), and in humans, the effects of nonglobal contexts (e.g., Dulsky, 1935; Murnane & Phelps, 1993, 1994, 1995; Weiss & Margolius, 1954; Wright & Shea, 1991), nonincidental contexts (e.g., Pan, 1926; Tulving & Thomson, 1973), and internal states (e.g., Blaney, 1986; Eich, 1995, 1989; Ucros, 1989) on memory. Research on transfer-appropriate processing (e.g., Blaxton, 1989; Kolers & Roediger, 1984; Morris, Bransford, & Franks, 1977) is also closely related to context-dependent memory. These effects are not in dispute, and they are not included in the present meta-analysis because the inclusion of such robust effects might cause us to overestimate the magnitude of incidental global context effects.

Two Guiding Principles of Incidental Environmental Context Encoding and Cuing

The body of published findings on incidental environmental context-dependent memory can be explained within a simple theory of incidental contextual processing. The guiding principles of the theory describe qualitatively how rules for contextual information processing apply in relation to various experimental paradigms that involve incidental context-dependent memory. Here the heuristic value of this theory will then be evaluated in a meta-analysis of published studies on incidental environmental context-dependent memory.

The two guiding principles that describe and explain effects of incidental environmental context on human memory are as follows: (1) The effect of environmental manipulation decreases as the use of nonenvironmental cues, at learning or at test, increases. For example, forming interitem associations at learning (which can overshadow or decrease learning of context), or using associative retrieval cues at test (which can outshine or overpower weaker context cues), should decrease the effect of environmental change by increasing dependence on interitem associations. (2) The effect of changing the environment from study to test decreases to the extent that subjects are encouraged at test to mentally reinstate the context that was present during learning (the mental reinstatement principle). Both of these principles indicate that the effects of environmental manipulations are reduced when attention is drawn away from the learning and/or test environments. This pattern can occur either because of a focus on associative information (overshadowing and outshining), or because the learning context is mentally reinstated (mental reinstatement).

Both of these basic principles can be extrapolated from the theory that memory derives from a need to maintain an ongoing mental representation of the environment (Glenberg, 1997). To cope with important environmental stimuli, perceptual systems are supported by memory systems, both sharing the same cognitive resources. To free up cognitive resources for thinking about something that is not in one's immediate environment (such as semantic associations), it is helpful to disengage from one's processing of the immediate environmental surroundings. Glenberg, Schroeder, and Robertson (1998) found, for example, that remembering was improved when subjects averted their gaze from their immediate experimental surroundings, a "remembering behavior" that facilitates disengagement from one's environment. Conceptual thought, according to this theory, is accomplished by an effortful suppression of one's processing of the ambient environment. Therefore, when memory retrieval is guided by interitem associations or by mental reinstatement of nonambient environmental cues (i.e., thinking about places other than one's current environment), such conceptually guided retrieval causes suppression of the environment in which retrieval takes place. This environmental suppression is the basis of overshadowing (if it occurs during learning) and outshining (if it occurs during retrieval), the nullification of incidental environmental cues caused by the use of noncontextual cues (see, e.g., Smith, 1988, 1994).

The degree of overlap or similarity between probes and memory traces determines memory. The more matching features there are in both a probe and a memory trace, the more likely it is that the memory trace will be retrieved. This similarity principle is consistent with the notion of encoding specificity, which states that episodic remembering can be cued by information encoded in the sought after memory trace (see, e.g., Tulving, 1983; Tulving & Thomson, 1973). This similarity principle is basic to many current theories of memory, such as CHARM (J. M. Eich, 1982), TODAM (Murdock, 1997), MINERVA 2 (Hintzman, 1988), SAM (Raaijmakers & Shiffrin, 1981), the General Context Model (Murnane & Phelps, 1994, 1995), and the Matrix model (Humphreys, Bain, & Pike, 1989). It should be clear, then, that memory for any particular episode can be enhanced by including in a memory probe the item, associative, and/or contextual information that is encoded in the memory trace corresponding to that episode.

The incorporation of any particular piece of information in either a memory trace or a probe depends on the policy that is current when traces and probes are constructed. In accordance with Glenberg's (1997) idea that the environment must be suppressed for conceptual processing to occur, we propose that one's default policy is to encode immediate environmental contextual features in memory traces and probes. People tend to be aware of their surroundings even when memorizing something. As such, environmental features are typically encoded along with the to-be-remembered material. Conditions at learning, however, might encourage subjects to focus more on the study materials to the exclusion of the environment, in which case environmental change will have less effect.

Given some effort, the environment can be suppressed to allow better conceptual processing (Glenberg, 1997; Glenberg et al., 1998). Thus, if subjects engage in extra conceptual processing during learning to encode inter-item associations, it may result in little or no encoding of the environmental context. When environmental suppression occurs at the time of learning, it can result in *overshadowing* (see, e.g., Matzel, Schachtman, & Miller, 1985), or a failure to store contextual information in memory. If contextual information is overshadowed at learning, then, according to the principle of encoding specificity, contextual cues provided at test will have no effect. *Outshining*, on the other hand (e.g., Smith, 1988, 1994), refers to the idea that the environment can be suppressed at test, diminishing the likelihood that ambient environmental information will be used in the construction of memory probes. Even if contextual information has been encoded successfully during learning, contextual cues at test may be ineffective if they are not used by the rememberer. The outshining hypothesis states that the use at test of noncontextual cues, such as cues that make use of inter-item associations, can diminish the subject's use of ambient contextual cues, thereby decreasing the influence of environmental manipulations (see Smith, 1994). We contend that both overshadowing and outshining can diminish the likelihood that environmental contextual manipulations will affect memory.

In addition to outshining and overshadowing that can be caused by environmental suppression concomitant with conceptual thinking, another reason that incidental environmental manipulations may not affect memory is that people can mentally reinstate nonambient environmental contexts, using imagined contextual features rather than the immediate perceptual context (e.g., Smith, 1979, 1984). To mentally reinstate a nonambient learning context, the subject might remember and image the learning environment. Mental reinstatement of nonambient contexts can be used as a mnemonic device employed at the time of learning, as is demonstrated by the method of loci. Mental reinstatement of an environmental context at the time of testing has also been shown to evoke memories of events originally experienced in the reinstated context (e.g., Smith, 1979, 1984).

Mental reinstatement of nonambient contexts causes suppression of the ambient context, limiting its encoding in memory traces and probes. The encoding of nonambient contexts leads to the storage of memory traces that do not encode experimentally manipulated environmental contexts. Furthermore, remembering can be independent of experimental environments if mental reinstatement is used by participants at test.

In summary, these basic theoretical principles provide a means of understanding the role of incidental environmental contexts as memory cues. The reinstatement of context cues, which are encoded by default in memory traces and probes, should benefit memory for information learned in the reinstated environment. This idea will be referred to as the *reinstatement hypothesis.* Conceptual processing at the time of learning can cause environmental suppression, limiting the degree of contextual encoding in critical memory traces, and thereby diminishing the effectiveness of reinstated context cues. This notion will be referred to as the *overshadowing hypothesis.* Environmental suppression can also occur at the time of a memory test because of conceptual processing, diminishing reinstatement effects even if context information is encoded in relevant memory traces. This refers to the *outshining hypothesis.* Finally, mental reinstatement of nonambient contexts at either learning or test can diminish the effects of ambient context cues, the *mental reinstatement hypothesis.* A meta-analysis of environmental context-dependent memory studies will now be reported as a test of these four basic hypotheses.

A Meta-Analysis

Briefly, meta-analytic techniques provide a way to quantify and combine findings across studies (Hedges & Olkin, 1985). With statistical techniques that are conceptually similar to analysis of variance, effect sizes calculated from individual studies can be used to provide a rich database that can augment a narrative review. The use of a meta-analytic technique is especially relevant for the present review. Context effects are often small and at times statistically nonsignificant. Meta-analysis can help reveal the conditions under which environmental change does not matter, as well as the conditions under which it does.

Hypotheses

Four primary hypotheses guided the structure of the meta-analysis: the reinstatement, outshining, overshadowing, and mental reinstatement hypotheses. Some predictions are the same for more than one of these hypotheses, whereas other predictions are critical tests of individual hypotheses. Another small set of hypotheses is subsidiary to the major theoretical hypotheses, but these are examined because of their potential relevance to questions about incidental environmental context-dependent memory.

Reinstatement. The reinstatement hypothesis states that across all studies, environmental context effects are reliable. In contrast, the null hypothesis suggests that incidental environmental context effects are not reliable (Fernandez & Glenberg, 1985).

Outshining. The outshining hypothesis (Smith, 1988, 1994; Smith & Vela, 1986) is based on the idea that when noncontextual cues are used to guide memory, the effects of contextual cues are diminished or eliminated. Therefore, less incidental context dependence is predicted for cued recall than for free recall, because the inclusion of noncontextual information (i.e., verbal cues) in memory probes should be greater in cued recall. Recognition tests typically provide even more noncontextual cues because the many words tested can reinstate associations and subjective organizations encoded during study sessions. Recognition tests should therefore show even less context dependence. Therefore, type of test should be shown to modulate environmental context dependence.

The outshining hypothesis also predicts associative input processing to lead to little or no effect of environmental context manipulations, whereas nonassociative input processing is predicted to lead to the largest effect of environmental context manipulations. If associative information is encoded and stored in memory traces, the same associative information can be self-generated and used to guide memory searches, even if physical cues are not provided at test.

Overshadowing. The overshadowing hypothesis states that if one's incidental environmental context is suppressed during learning, then environmental information will not be encoded and stored in memory, thereby reducing or eliminating effects of experimenter-manipulated environments on memory. For example, if associative processing at input causes suppression of the environment, little or no environmental context will be encoded and stored in episodic memories, which is consistent with the overshadowing hypothesis. The overshadowing hypothesis shares this prediction with the outshining hypothesis, albeit for different reasons.

A prediction made only by the overshadowing hypothesis, however, is that time-in-context at input should be positively related to the finding of environmental context-dependent memory. The longer the time one has spent in an environment, the greater should be the chances of encoding and storing environmental information in memory traces. The outshining hypothesis makes no such prediction.

Mental reinstatement. If learning contexts are mentally reinstated by subjects even when they are tested in new environments, then context effects should be diminished or eliminated. Therefore, instructions to mentally reinstate the learning environment should reliably improve memory for subjects tested in unfamiliar contexts.

The mental reinstatement hypothesis also predicts that context dependence should be modulated by the type of test used. For one thing, it has been suggested that the memory cues provided by recognition tests may encourage subjects to mentally reinstate the learning context (see, e.g., Bjork & Richardson-Klavehn, 1988). The limited cues provided by free recall tests, on the other hand, may be less likely to encourage mental contextual reinstatement.

Bjork and Richardson-Klavehn (1988) also distinguished between *first-order* and *second-order* experimental paradigms, defining first-order paradigms as those in which contexts are either matched or mismatched at test. Second-order paradigms use multiple contexts and multiple lists, such as the interference reduction paradigm (see, e.g., Bilodeau & Schlosberg, 1951; Greenspoon & Ranyard, 1957). Savvy subjects may think to mentally reinstate their learning contexts in first-order (reinstatement) paradigms as a way of improving recall performance, thus diminishing observed context effects. In second-order paradigms, however, mental reinstatement of contexts gives subjects no performance advantage; reinstatement of interfering contexts in interference reduction paradigms can only hinder recall, and mental reinstatement of multiple input contexts yields no memory advantage (Smith, 1979, 1984). Because second-order paradigms do not encourage subjects to suppress the ambient test environment in the way that first order paradigms do, it was predicted that the typical (first-order) reinstatement paradigm should be less effective in producing environmental context effects than interference reduction or multiple context paradigms.

Rationale of Meta-Analysis

In the following meta-analysis, we examine the reliability of environmental context-dependent memory findings and identify the factors that appear to influence such effects. The techniques of meta-analysis used in the present review include those described by Hedges and Olkin (1985). The techniques allow one to estimate the *effect size* of a treatment variable across an entire body of research findings. To provide a unit of analysis comparable across studies, a standardized effect size (g) is computed for each study by subtracting the experimental and control group means (e.g., same- and different-context groups) and dividing by the standard deviation assumed to be common to the two conditions. A positive effect size indicates that the experimental group mean is larger than the control group mean, and a negative effect size indicates the opposite. The effect size magnitude is an indication of the standardized difference magnitude between the means. This technique, however, provides a biased estimate of the effect size (slightly higher than the "true" effect size) that increases as the number of participants used in a given study decreases. A corrected effect size (d) was therefore calculated for each effect size (Hedges & Olkin, 1985).

The techniques of meta-analysis are of particular relevance given the ambiguity concerning the parameters of context dependency. Studies in which statistically significant effects of context reinstatement have not been found may nevertheless combine to reveal effect sizes significantly different from zero. In addition, variables that modulate the treatment in question can be identified, and evaluations can be made as to whether or not such groupings derive from common populations.

Criteria for Inclusion in the Meta-Analysis

Four main criteria guided the inclusion of studies in the present meta-analysis. First, inclusion was limited to studies in which context was at least intended to be incidental to target acquisition. In one of Eich's (1985) conditions, for instance, participants were instructed to encode target items as existing at specific loci in their environmental context. In the other main condition, no such instructions were given. The former condition was therefore not included in the present meta-analysis, whereas the latter condition was. It should be noted that although the typical context manipulation involves the room in which target acquisition and testing occurs, other forms of context were also included in the analysis. These include, for example, Godden and Baddeley's (1975) underwater versus surface manipulation and Dallett and Wilcox's (1968) "box environment," in which participants did not change their room location but rather changed the box environment, fitted over their heads.

Another criterion for inclusion in the meta-analysis is that context be "global" or slow changing (Glenberg, 1979). Generally, this involves the acquisition of target information that occurs in a single environmental context. Paradigms that utilize multiple learning environments were also included because they involve the acquisition of entire bodies of list items learned in a single environmental context, even though multiple lists may each be learned in a separate environmental context. An example of context manipulation that does not meet this criterion was reported by Weiss and Margolius (1954). In that experiment, paired associates were presented on different-colored cards. Later relearning was assessed under one of various background color manipulations. The key factor here is that the background context of card color changed for each paired associate. The background color was a fast changing and therefore nonglobal form of context manipulation. Studies with similar fast changing context manipulations were also excluded from the analysis (e.g., Dulsky, 1935; Murnane & Phelps, 1993, 1994, 1995; Wright & Shea, 1991).

Third, drawing on the distinction between extrinsic and intrinsic forms of context (Hewitt, cited in Godden & Baddeley, 1975), we have included in the analysis only contexts that may be considered to be "external" or extrinsic to target items. The spirit of the distinction may be captured in such studies whose context manipulations include speaker's voice (Geiselman & Bjork, 1980; Geiselman & Glenny, 1977) or the auditory versus visual manipulation of input and test list items (e.g., Pessin, 1932). These studies have manipulated context that may be considered "intrinsic" or part of the stimulus items themselves.

The meta-analysis also excludes studies with nonhuman subjects. There is no question about the robust findings of context-dependent learning and memory in animals. The inclusion of such studies would "stack the deck" in terms of finding overall context-dependent effects. Nonhuman studies would be difficult to interpret in terms of the often very different paradigms used in human and nonhuman experiments. Finally, questions about conceptual processing and suppression of the environment do not clearly apply to nonhumans, who may be more environmentally bound than humans.

Other Studies Not Included

As previously discussed, studies in which researchers manipulated internal mood or pharmacological state (e.g., Bower & Cohen, 1982) or physical/physiological state (e.g., Rand & Wapner, 1967, manipulated body posture during input and test) were not included in the analysis. Although these studies are relevant, such factors do not fall within what is considered to be clearly extrinsic forms of context.

Three studies conducted by Block (1982, Experiments 1–3) were also not included because the dependent measures were not clearly comparable with others included in the meta-analysis. The studies reported by Block showed that context change can reduce "positive-time-order-error": The tendency to overestimate the first of two consecutive time durations was reduced when a context shift accompanied the second time duration. The effect sizes associated with each of Block's experiments ranged from moderate to large.

Studies in which researchers manipulated olfactory context (e.g., Cann & Ross, 1989) were also not included in the meta-analysis. Although olfactory context does meet the requirements for inclusion, the fact that so very few studies of olfactory context exist, and that the modality shift from physical (implicitly operationalized as the visual context) to olfactary modality is sensorily distinct, led us to simply not consider this form of context manipulation in the meta-analysis. The paucity of research concerning indirect tests of context-dependent memory also affected our decision of what not to include in the meta-analysis. The only published study with an indirect test of memory was reported by Smith, Heath, and Vela (1990). In that study, priming of the lower frequency spellings of homophones was greater in the context in which the lower frequency spellings of homophones were originally encoded.

Finally, classroom reinstatement studies were excluded from the meta-analysis. Although manipulation of classroom context can provide an ecologically valid assessment of context change on memory performance, it poses problems of interpretation, owing to the dependent and independent variables used in classroom studies. Unfavorable characteristics of many classroom studies include the lack of consistent and clear measures of memory, as well as uncontrolled extra-classroom study contexts.

Many of these reported studies do not specify the type of memory test used (e.g., essay, true/false, matching) but rather report only test grade differences or error rate differences. Furthermore, and most damaging, study contexts are generally not controlled, so that students are free to study in contexts other than their classroom environment. Memory can be increased when multiple study contexts are used (see, e.g., Smith, 1982, Experiments 1 and 3, 1985), and context-dependent recall can be nullified when the acquisition of list items occurs in many contexts (Smith, 1982, Experiment 3). Assessment of classroom context manipulations is thus compromised when extra-classroom study contexts are not controlled.

In summary, the main criteria for inclusion in the analysis involved context that was (1) incidental to target acquisition, (2) global and associated with an entire body of target items, (3) extrinsic to both the target items and to the subjects, and (4) manipulated and tested on human subjects.

METHOD

A total of 93 separate effect sizes were calculated from 75 studies reported in 41 separate articles conducted between 1935 and 1997. The studies reviewed include those identified from a computer search and from a manual survey of *Psychological Abstracts* (1900–1997) from reference sections of published studies. The keywords used in the computer search included *context*, *environment*, and *memory*.

Listing and Discussion of Coded Variables

Paradigm. The first major methodological concern that guided variable selection involved the first- and second-order context effect distinction suggested by Bjork and Richardson-Klavehn (1988). The typical first-order environmental context paradigm was coded as a reinstatement effect (i.e., a memory test that occurs in an environmental context that is the same as, or different from, that for input). The second-order paradigm was subdivided into two separate categories. One category included interference reduction studies that measured retroactive or proactive interference reduction when lists were learned in separate environments, as compared with when they were learned in the same environment. The remaining category included multiple input room studies in which memory for a single list or multiple lists was measured when learning occurred in single versus multiple environmental contexts. The three levels were coded as reinstatement effects, interference reduction effects, and multiple input room effects.

Test type. The three direct memory test types coded were recall, recognition, and cued recall.

Encoding orientation. The orienting task at input was judged according to the degree of interitem associative processing engendered by task instructions and was coded as associative processing, nonassociative processing, or "unsure." Examples of associative interitem processing included procedures that utilized sorting tasks, paired associates, or instructions that directed participants to relate list items in meaningful ways. Studies that directed the participant to engage in primary or rote rehearsal, pleasantness ratings, and other forms of item-specific processing were coded as nonassociative processing procedures. Finally, the "unsure" category included procedures that did not direct the participant to engage in a specific orienting task, or when the data reported did not allow assessment of effect sizes in terms of separate orienting task instructions (e.g., Cousins & Hanley, 1996).

Study type. The two levels of this variable were coded as either physical reinstatement or imaginal reinstatement. The division of analyses into these separate subsections was deemed appropriate for a number of reasons, all of which concerned the need for clarity. Each section of the meta-analysis examines different aspects of context manipulations that, within each grouping, cohere both conceptually and methodologically. The effect sizes included in the first section (physical reinstatement studies) provide a database for studies in which physical context is manipulated and direct measures of memory are reported (e.g., recall, recognition, cued recall). Questions concerning the reliability of environmental context reinstatement and possible modulating variables can be most clearly addressed in this section.

Whereas the section involving mental reinstatement is also based on direct measures of memory, the method of context reinstatement is methodologically distinct from physical reinstatement. In addition, several of the effect sizes derived from eyewitness identification studies involving the recognition of a face in a lineup as opposed to recall or recognition of word lists. Several of these studies also employed mental reinstatement techniques (e.g., Krafka & Penrod, 1985; Malpass & Devine, 1981) that were much more elaborate and systematic than simple mental reinstatement instructions (e.g., Smith & Vela, 1992). Furthermore, imaginal and environmental context manipulations were confounded in some studies (e.g., Krafka & Penrod, 1985; McSpadden, Schooler, & Loftus, 1988).

Experimenter at different-context condition. Changing the experimenter for the different-context condition could remove a source of possible context reinstatement. It could also control for possible experimenter bias effects. The levels of this variable were coded as either same experimenter as input, or different experimenter from input.

Mode of stimulus presentation. It is possible that the mode of stimulus presentation influences the degree to which context information is initially encoded. With visual stimulus presentation (e.g., slides, booklets, computer screens), participants may be induced to pay relatively less attention to their ambient context than if target information is presented aurally. With aural presentation, participants may be more likely to scan their ambient environment and thereby encode more context information. Levels of this variable were coded as visual presentation or aural presentation of target information.

Number of input items. This variable was restricted to list learning studies. Memory for a relatively large number of target items may be differentially sensitive to context manipulations in relation to memory for a relatively small number of target items. This variable was coded along a continuous scale.

Duration of stimulus exposure. Also coded along a continuous scale, longer exposure time may increase the amount of processing devoted to each stimulus. More processing time may lead to more stable and/or accessible memories and hence less dependence on context information.

Duration of total input context exposure. It is possible that context dependence will vary as a function of the length of time spent in the input context. A significant positive effect of this variable would provide support for the overshadowing hypothesis. This variable was also coded along a continuous scale.

Input–test interval. It is possible that the input–test retention interval modulates context dependency. Fernandez and Glenberg (1985), for example, noted that many anecdotal reports of context dependency involve long retention intervals. The four levels of this variable were coded as less than 5 min, 5 min to 1 h, 1 h to 1 day, and 1 day to 1 week.

Method of Analysis

Following Hedges and Olkin (1985, pp. 147–165), the calculations and procedures for fitting models involving categorical variables were used. These procedures have an advantage over the more traditional analysis of variance by providing a within-group fit statistic (Q_W) that assesses whether studies within groups are homogenous. The Q_W test statistic has an approximate chi-square distribution with $n - 1$ degrees of freedom, where n is the number of studies included within a grouping. A significant Q_W indicates that substantial heterogeneity remains in the model and that more fine-grained analysis is necessary.

Hedges and Olkin (1985) provide a systematic approach to categorical model fitting. The first step in this approach is to determine whether the effect sizes for all studies are homogeneous (Q_T). Given a significant deviation from homogeneity, the next step is to break down the studies into important a priori groupings (e.g., first- and second-order context effects), resulting in what essentially is a blocking variable with two or more levels. A Q_B and Q_W test statistic is then computed to test, respectively, for between- and within-group homogeneity. Each statistic has an approximate chi-square distribution with $p - 1$ (Q_B: where p is the number of between-groups) and $m - 1$ (Q_W: where m is the total number of effect sizes within groups) degrees of freedom. The Q_B goodness of fit statistic is a direct analogue of the analysis of variance main effect. A significant Q_B indicates that the average weighted d differs across groups. The Q_W goodness of fit statistic provides a test of group homogeneity. A significant Q_W rejects the null hypothesis that group homogeneity exists and suggests that the levels of the first blocking factor must be divided further, using another categorical variable. The process of subdividing and testing for the between- and within-group fit stops when within-group homogeneity is reached.

Variables coded along a continuous scale served as predictors on which effect size estimates were regressed. The effects of these variables were evaluated with a least squares regression procedure detailed by Hedges and Olkin (1985, pp. 167–188). Each regression provides a test for the effect of the predictor variable (transformed to a z score) and a test of the model specification (Q_E). The Q_E error-sums-of-squares statistic has an approximate chi-square distribution with $k - p - 1$ degrees of freedom, where k is the number of effect sizes and p is the number of predictors.

RESULTS AND DISCUSSION

A summary of the experiments included in the meta-analysis is provided in the Appendix. Across all studies, the average weighted effect size was significant ($d = .28$; 95% CI + .05, $n = 93$), supporting the hypothesis that environmental context effects are reliable. As may be expected, the hypothesis that the studies were drawn from a common population was rejected, with the goodness of fit statistic revealing substantial heterogeneity (i.e., significant differences) among the effect sizes ($Q_T = 242.47$, $p < .001$). What this implies, as we will see below, is that under some circumstances the effect is small and in other circumstances it is quite a bit larger.

The remaining analyses are reported in two major sections, with each section pertaining to a different subset of studies. Analyses of 83 effect sizes calculated from studies (89% of the entire sample) that did not include imaginal reinstatement instructions are reported in the first section. The second section reports the results for 10 effect sizes from the remaining studies in which a mental reinstatement procedure was used.

Physical Reinstatement Studies

The average weighted effect size for all the studies included in this section was significantly greater than zero ($d = .28$; 95% CI + .02, $n = 83$). The effect sizes also deviated significantly from homogeneity ($Q_T = 193.22$, $p < .001$).

Main Effect Analyses

A summary of the main effect analyses is shown in Table 1. Each discretely coded variable was subjected to an analogue of the analysis of variance test for main effects. As stated earlier, the Q_B homogeneity statistic tests whether the levels of a variable are significantly different from each other. Given that substantial heterogeneity may remain within the levels of a variable, however, cau-

Table 1
Summary of Main Effect Analyses for Physical Reinstatement Studies

Variable	Levels	Weighted Effect Size (d)	95% Confidence Interval for d	
			Lower	Upper
Paradigm*	Reinstatement	.23	.17	.29
	Interference reduction	.68	.47	.89
	Multiple contexts	.45	.26	.64
Test type	Recall	.29	.21	.37
	Recognition	.27	.18	.36
	Cued recall	.25	−.02	.52
Processing at Input*	Associative	.13	.03	.23
	Nonassociative	.33	.20	.46
	Unsure	.38	.29	.47
Mode of stimulus presentation	Auditory	.28	.19	.37
	Visual	.17	.08	.26
Experimenter at DC*	Same as input	.26	.12	.40
	Different from input	.62	.48	.76
Learn/test interval*	Less than 5 min	.27	.17	.37
	5 min to 1 h	.22	.13	.31
	1 h to 1 day	.28	.15	.41
	1 day to 1 week	.63	.42	.84
Study type†	Physical reinstatement	.28	.26	.30
	Imaginal reinstatement	.26	.13	.39

*Differences among effect sizes were statistically significant. †These data are included in the summary in order to highlight the similar effect sizes for both types of studies.

tion must be made when one is interpreting the Q_B statistic. The conclusions of the main effect analyses of moderating variables were supported by a more fine-grained model-fitting approach (Hedges & Olkin, 1985).

The effect of *paradigm* (reinstatement, interference reduction, and multiple context effects) was significant ($Q_B = 18.60$, $p < .001$), indicating that the average weighted effect sizes differed across groups. The effect size for reinstatement studies was $d = .23$ (95% CI + .06, $n = 66$), for interference reduction studies $d = .68$ (95% CI + .21, $n = 10$), and for multiple-context studies $d = .45$ (95% CI + .19, $n = 7$). A priori contrasts (Hedges & Olkin, 1985) revealed a significant difference in the average weighted effect size between reinstatement and interference reduction studies ($\chi^2 = 15.47, p < .001$), and a significant difference between reinstatement and multiple-context studies ($\chi^2 = 4.55$, $p < .05$). Partitioning the within-group fit statistic Q_W according to the levels of paradigm (Q_{wi}) revealed within-group heterogeneity for both reinstatement and interference reduction studies (respectively, $Q_{w1} = 140.21$, $p < .001$, and $Q_{w2} = 28.16$, $p < .001$), but did not yield within-group heterogeneity for multiple context studies ($Q_{w3} = 6.25, p > .05$).

The effect of *test type* (recall, recognition, and cued recall) was not significant ($Q_B = .21, p > .05$). The averaged weighted effect size for free recall was $d = .29$ (95% CI + .08, $n = 52$); for recognition, $d = .27$ (95% CI + .09, $n = 28$); and for cued recall, $d = .25$ (95% CI + .27, $n = 3$). Partitioning the within-group fit statistic into that associated with each level of test type showed only that studies using cued recall were not heterogeneous ($Q_{w1} = 1.62$, p > .05), suggesting that free recall and recognition studies required further subdivision. However, when compared across all studies, free recall effect sizes were not reliably larger than those associated with recognition and cued recall (though both recall and recognition effect sizes were significantly larger than zero).

The effect of *encoding orientation* (associative processing, nonassociative processing, and unsure) was significant ($Q_B = 15.37, p < .001$). The average weighted effect size for studies coded as engendering associative inter-item processing at input was $d = .13$ (95% CI+ .10, $n = 29$); and for nonassociative processing at input, $d = .33$ (95% CI + .13, $n = 12$). Studies in which input processing was coded as unsure revealed the largest effect size, $d = .38$ (95% CI + .09, $n = 42$). An a priori test comparing the average associative and nonassociative effect sizes was significant ($\chi^2 = 6.67, p < .01$), indicating that across studies the average nonassociative effect size was larger than the average associative effect size. An a posteriori test showed that the average nonassociative effect size did not differ from the average unsure effect size ($\chi^2 = .42$, $p > .05$). The partitioned Q_{wi} fit statistic revealed a significant effect only for the unsure processing category ($Q_w = 128.49, p < .001$).

The effect of of *mode of stimulus presentation* (visual vs. aural) was not significant ($Q_B = 2.68$, $p > .05$), although a larger effect size was associated with aural presentation ($d = .28$, 95% CI + .09, $n = 37$) than with visual presentation ($d = .17$, 95% CI + .09, $n = 37$). The partitioned within-group fit was significant for both visual presentation ($Q_{w1} = 86.92$, $p < .001$) and auditory presentation ($Q_{w2} = 49.13$, $p < .02$).

The effect of *experimenter at different context condition* (same as input context vs. different from input context) was significant ($Q_B = 41.93$, $p < .001$). The presence of the input-context-experimenter at the different context condition ($d = .26$, 95% CI + .14, $n = 60$) resulted in a lower average effect size than when the experimenter was changed ($d = .62$, 95% CI + .14, $n = 18$) for the different-context condition. The partitioned within-group fit was significant for each level indicating substantial heterogeneity remained in each group (same experimenter as input context, $Q_{w1} = 93.66$, $p < .001$, and different experimenter from input context, $Q_{w2} = 43.30$, $p < .01$).

The effect of *input–test interval* (less than 5 min, 5 min to 1 h, 1 h to 1 day, and 1 day to 1 week) was significant ($Q_B = 12.31$, $p < .01$). The average effect sizes associated with each level were as follows: less than 5 min, $d = .27$ (95% CI ± .10, $n = 26$); 5 min to 1 h, $d = .22$ (95% CI ± .09, $n = 30$); 1 h to 1 day, $d = .28$ (95% CI ± .13, $n = 21$); and 1 day to 1 week, $d = .63$ (95% CI ± .21, $n = 6$). An a posteriori test comparing the longest retention interval with the average of the remaining retention intervals was significant [$\chi^2(1) = 6.89$, $p < .02$], suggesting that the longest retention interval coded produced an effect size larger than the remaining shorter retention intervals. Partitioning the within-group fit revealed heterogeneity for each level (respectively, $Q_{w1} = 62.07$, $p < .001$; $Q_{w2} = 63.79$, $p < .001$; $Q_{w3} = 43.58$, $p < .005$; $Q_{w4} = 11.54$, $p < .05$).

Regression Analyses

Variables coded along a continuous scale served as predictors on which effect size estimates were regressed. These variables included the number of input items to be learned and the duration of total input context exposure (in seconds). The effect of these variables was evaluated separately with the least squares regression procedure detailed by Hedges and Olkin (1985). Each regression provided a test for the effect of the predictor variable (transformed to a z score) and a test of the model specification (Q_E).

The natural log of the number of input items (ranging from 1 to 240) was not a significant predictor of effect size ($z = -.84$, $p > .05$). Similarly, the natural log of the total time in the learning context was not a significant predictor of effect size ($z = -.54$, $p > .05$). Finally, the natural log of the duration of stimulus exposure also did not predict effect size ($z = -.93$, $p > .05$).

A Descriptive Model

The method for fitting categorical models (Hedges & Olkin, 1985) was used to provide a more coherent and fine-grained analysis of context effects in memory. It should be noted that the term "model" is simply a quasi-hierarchical summary of the variables included in the analysis. Beginning at the top of the hierarchy, categorical groupings were tested for heterogeneity until groupings were found that showed no further significant heterogeneity. The model-fitting analyses corroborated all of the main effects analyses; therefore, we have presented only a brief summary of the variables included in the model and the results of the analyses in Table 2.

GENERAL DISCUSSION

The results of our meta-analysis provide varying degrees of support for each of the four major hypotheses previously described—the reinstatement, outshining, overshadowing, and mental reinstatement hypotheses. Averaged across all studies, the meta-analysis indicates that manipulations of incidental environmental contexts have reliably affected memory. The average weighted effect

Table 2
A Descriptive Model of Context Reinstatement Studies

I. First-Order /Context Reinstatement ($n = 66$) ($d = .23$; 95% CI +.06)						
Recall ($n = 37$)			Recognition ($n = 27$)			*Cued Recall ($n = 2$)
.29 (+.08)			.27 (+.09)			.25 (+.27)
Associative Processing			Associative Processing			
Yes*	Unsure	No[a]	Yes*	Unsure[a]	No[a]	
.01	.32	.33	.18	.28	.33	
(±.14)	(±.18)	(±.22)	(±.19)	(±.15)	(±.16)	
II. Second-Order/Interference Reduction ($n = 10$) ($d = .68$; 95% CI +.21)						
Recall ($n = 9$)			Recognition[b]			Cued-Recall ($n = 1$)
.70 (+.23)						
Associative Processing						
Yes*	Unsure	No[b]				
.22	1.23					
(±.32)	(+.34)					
III. *Second-Order/Multiple Context ($n = 7$) ($d = .45$; 95% CI ±.19)						

Note—Numbers reported under each heading include the effect size and respective 95% confidence interval. *Indicates within-group homogeneity. [a]The experimenter at the different context variable accounted for the heterogeneity within this level. [b]There were no studies that fell within this grouping.

size across all studies indicates that context has a modest (d = .28) but reliable (95% CI + .05) effect on memory performance. This finding indicates that changing rooms and other similar forms of incidental environmental context affects memory for events experienced within those environments. This conclusion appears to contradict that of Fernandez and Glenberg (1985), the title of whose paper states, "Changing environmental context does not reliably affect memory." Although on the average, reported environmental context-dependent memory effect sizes are reliable, a glance at the set of all published results (Appendix) shows that changing rooms is not sufficient for causing context effects when circumstances are not favorable. For example, effect sizes were rarely greater than zero in the 28 studies that used associative processing at input, and in those few, the effect only barely exceeded zero. Clearly, when associative processing is used at input, as was true in 7 of the 8 cases reported by Fernandez and Glenberg, changing rooms does not appear to affect memory.

It is also worth noting that context effects are not always modest. Effect sizes larger than d = .80 are considered to be large effects (Cohen, 1977). Of the 93 effect sizes considered in the present meta-analysis, 18 (nearly 20%) were greater than .80. For example, using an interference reduction paradigm, Greenspoon and Ranyard (1957) reported a very large effect size of d = 1.75, and Dallett and Wilcox (1968) found an effect size of d = 1.48. In other examples, using an imaginal reinstatement paradigm, Krafka and Penrod (1985) found a large effect size of d = 1.71, and Smith (1979) demonstrated an effect of d = .88. Thus, the modest average effect sizes reported in the present meta-analysis should not obscure the fact that in many cases large effects have been found.

Switching experimenter in the different-context condition was associated with increased context reinstatement effect sizes. From a methodological perspective, this variable appears to contribute importantly to context dependency. The interval of time between input and test was also related to context dependency; the longest retention interval (1 day to 1 week) was associated with an effect size significantly larger than that found for shorter retention intervals. Anecdotal reports of context reinstatement often involve rather long retention intervals, such as the apparent context cuing that many people experience at school reunions. Our finding with respect to retention interval seems consistent with such anecdotal reports, but however suggestive, it does not allow us to directly evaluate the importance of very long retention intervals in context-dependent memory.

The results of the meta-analysis provided mixed support for the outshining hypothesis. The outshining hypothesis was supported by the finding that the type of orienting task at input is an important modulator of the magnitude of context effects (for direct tests of memory). Nonassociative processing at input was shown to produce an average effect size (d = .33) significantly larger than that engendered with associative interitem processing at input (d = .13). If target stimuli are encoded and stored in memory in association with meaningful materials, subjects are likely to generate those associations from memory and suppress the ambient test environment, using associative information to guide retrieval. Without associative interitem processing at input, it is unlikely that retrieval will be guided by associative cues, and therefore more likely that manipulated environmental cues will be effective memory aids.

Surprisingly, there was not a significant effect of the type of memory test (i.e., free recall, cued recall, or recognition), contrary to a prediction of the outshining hypothesis. This lack of an effect is surprising, because experimental comparisons of recall and recognition have clearly shown environmental manipulations that affected recall performance, but not recognition memory (e.g., Godden & Baddeley, 1975, 1980; Smith et al., 1978). Averaged across all studies, however, recall and recognition had effect sizes greater than zero, whereas the average effect size associated with cued recall was not significantly greater than zero. The effect size for free recall (d = .29) was numerically only slightly larger than that for recognition (d = .27).

The lack of a significant effect of test type in the meta-analysis highlights a difference in meta-analytic as opposed to experimental tests of hypotheses. Why do the results of experimental comparisons appear to contradict the meta-analytic finding? One possibility is suggested by a finding of the meta-analysis related to associative processing at input; encoding interitem associations appears to dampen or eliminate the effects of environmental context manipulations. The direct experimental comparisons of recall and recognition (Godden & Baddeley, 1975, 1980; Smith et al., 1978) all used lists of words, which are highly associable, whereas many of the studies resulting in large context-dependent recognition effects (e.g., Dalton, 1993; Krafka & Penrod, 1985; Malpass & Devine, 1981; Smith & Vela, 1992) used less associable materials, such as faces of incidentally encountered confederates. The pattern that emerges from consideration of both experimental and meta-analytic results is that context dependence can be observed in recognition if other influential factors, such as nonassociative processing, are optimally orchestrated to yield context effects (e.g., Dalton, 1993; Smith, 1986). When those influential factors are controlled (e.g., when lists of words are used for both recall and recognition tests), free recall produces greater context dependence than does recognition.

Support for the overshadowing hypothesis as an explanation of incidental environmental context-dependent memory was mixed. Consistent with the overshadowing hypothesis was the finding that associative processing at input diminished context dependence. The overshadowing explanation of this result is that environmental information is suppressed at input, making memory of target stimuli less context dependent because less contextual information is encoded. Inconsistent with the overshadowing hypothesis was the finding that time in context did not predict the magnitude of the context effects. Briefer exposures to environments at learning were expected to decrease the chance that such contexts would be encoded and stored in memory. These findings do not reject the overshadowing hypothesis per se, a phenomenon that is well documented elsewhere (e.g., Mazur, 1986), nor do they contradict findings that attention to environmental information at input can increase memory's dependence on environmental cues (see, e.g., Eich, 1985).

The mental reinstatement hypothesis is clearly supported by the findings of the meta-analysis. In particular, the two results that support mental reinstatement are the finding that mental reinstatement instructions improved memory, and that first-order reinstatement paradigms were less likely to yield context dependence than were second-order paradigms. The latter result can be explained by the notion that subjects in first-order paradigms are motivated to mentally reinstate their input contexts, whereas those in second-order paradigms are not so motivated. Thus, mental reinstatement may make first-order paradigms less sensitive to experimentally manipulated environments.

Imagery reinstatement techniques varied across studies. For example, Malpass and Devine (1981) used a technique called "guided memory," in which participants at test were queried about their feelings and reactions to the input incident (a staged vandalism) and their memory for details of the room in which the incident had occurred. In addition to mental reinstatement, Krafka and Penrod (1985) provided store clerks with physical cues concerning the identity of the to-be-remembered "customer" (e.g., a nonphoto identification used by the confederate), confounding mental and physical reinstatement. The effect size associated with the Krafka and Penrod study was the largest among all effect sizes included in the analysis ($d = 1.71$).

The list-learning studies that manipulated mental reinstatement and were not confounded with physical reinstatement were reported by Smith (1979, 1984). The mental reinstatement technique used was relatively brief and involved having participants visualize the input room while they were in the different context, and list objects found in the room. Those studies reported moderate to large effect sizes (respectively, ds = .27 and .88).

The effect of paradigm also supported the mental reinstatement hypothesis. Interference reduction and multiple context effect paradigms were shown to produce large and reliable effect sizes. According to Cohen (1977), small effect sizes are on the order of $d = .20$ or less, medium effect sizes are in the range of $d = .50$, and large effect sizes are considered to be $d = .80$ and greater. By far, the largest average effect size derived from the subset of studies in which an interference reduction paradigm was used ($d = .68$). Multiple-context studies showed the second largest average effect size ($d = .45$), and reinstatement paradigms the smallest ($d = .23$). These findings are consistent with the assumption that participants can and do mentally reinstate their input context with reinstatement paradigms, serving to dilute the effects of physical manipulations.

Mood Mediation Hypothesis

The mood mediation hypothesis (see, e.g., Eich, 1995) states that findings of place- (environmental context) dependent memory may actually be caused by mood-dependent memory. The idea is that moods are hypothesized to be associated with memories, and that moods can later cue those associated memories. Furthermore, different environments can induce different moods, and the environmental changes that change moods the most should be most likely to result in place-dependent memory, a prediction consistent with Eich's (1995) results. Unfortunately, the published studies reviewed by our meta-analysis did not report subjects' moods, nor did they report how systematically the manipulated contexts might have altered moods. It is not clear that the mood mediation hypothesis could explain other findings of the meta-analysis, such as the importance of paradigm, test type, or stimulus modality. Therefore, mood mediation does not serve as an adequate alternative explanation of our findings.

Although the mood mediation hypothesis ascribes environmental context effects to very different causes than does our theory, Eich's (1995) result, that context shifts affected memory only when there were concomitant mood shifts, can be resolved to some degree by the mental context hypothesis (Smith, 1995). This hypothesis, derived from ideas expressed by McGeoch (1942), Bower (1972), and others, states that mood information can be incorporated into memory traces and probes, just as item, associative, and environmental information can be represented.

If we further assume that internal states, such as moods or drug-induced states, are more difficult to suppress than ambient environments, then subjects could overcome an environmental shift by suppressing the test environment and mentally reinstating the learning environment, but they could not suppress mismatched test moods in the same way. This notion could explain why, in Eich's (1995) study, mood shifts decreased memory whether or not environments were also shifted. It might be possible to test this explanation of Eich's results with indirect methods, such as second-order paradigms and indirect memory measures, to decrease the extent to which subjects are encouraged to mentally reinstate their learning contexts.

Environmental Context Encoding and Cuing

The theoretical principles we have identified provide a simple, but useful structure for describing incidental environmental context-dependent memory. The theory states that environmental contextual features are processed at learning and at test unless the ambient environment is suppressed, either because of conceptual processing, or because nonambient contexts are mentally reinstated. The theory is predicated on the idea that remembering depends on the similarity or featural overlap of memory traces and memory probes, a commonly used principle of memory theories (e.g., J. M. Eich, 1982; Hintzman, 1988; Humphreys et al., 1989; Murnane & Phelps, 1995; Raaijmakers & Shiffrin, 1981). The basic principles proposed herein can be easily accommodated by most current theories of memory and are consistent with the findings of the present meta-analysis.

One problematic aspect of the theory, however, concerns the way in which it explains the effects of environmental context manipulations on recognition memory. The theory predicts that the item information provided by recognition tests should displace contextual information in memory probes, thereby reducing the effects of environmental context on recognition tests. Indeed, the meta-analysis showed that of the direct memory measures, free recall tests were the most sensitive to environmental manipulations. Nonetheless, studies done with recognition tests also showed reliable context dependence, once their effect sizes were examined meta-analytically.

Furthermore, a number of experiments showing context-dependent recognition memory have been reported by Murnane and Phelps (1993, 1994, 1995). In these studies, contexts have been manipulated via configurations of stimuli on the computer screen where target words were learned and tested. In most of these experiments, it was found that reinstated screen contexts, relative to altered contexts, produce both higher hit rates and higher false alarm rates, referred to as "same-direction" effects. These findings contrast with the finding of our meta-analysis that shows that incidental environmental context manipulations reliably affect recognition memory (i.e., measures such as d'), rather than cause the same-direction effects found by Murnane and Phelps.

In spite of the differences between Murnane and Phelps's studies and those reviewed in the present meta-analysis, at least one finding appears to link the two. This finding is that, contrary to predictions of our theory, shallower input processing did not increase context-dependent recognition (Murnane & Phelps, 1995; Smith, Vela, & Williamson, 1988). Our theory predicts that deeper processing of items at input should increase the encoding of item information in memory traces and in memory probes, thereby decreasing context-dependent effects. Although decreasing item information by shallow processing does not increase context-dependent recognition, decreased interitem associations at input does increase context-dependent recognition (Smith, 1986). Thus, our theory is successful in predicting outshining effects based on the use of associative information in memory probes, but it is not as successful in its prediction of outshining due to the use of item information in probes.

Conclusions

In spite of some failed attempts to find environmental context-dependent memory effects, it is clear from our meta-analysis that across all reported studies, the effects are reliably found. The meta-analytic method is used here to provide quantitative analyses to augment our review of context-dependent memory findings. As such, neither the meta-analytic method nor any review technique can replace the experimental method in establishing cause and effect. What the meta-analytic review can provide is a systematic post hoc analysis of what may sometimes seem a confusing archive of results. The value of such an analysis is to then predict when environmental context-dependent memory effects will and will not be observed in subsequent experimental studies.

The results of our review and meta-analysis are consistent with the prediction that factors that encourage suppression of one's immediate environment should diminish the observed effects of manipulations of incidental environmental contexts. Such factors include associative processing at input, mental reinstatement of nonambient contexts, and the use of reinstatement paradigms, rather than second-order memory paradigms.

REFERENCES

Balsam, P. D., & Tomie, A. (Eds.) (1985). *Context and learning*. Hillsdale, NJ: Erlbaum.

Bilodeau, I. M., & Schlosberg, H. (1951). Similarity in stimulating conditions as a variable in retroactive inhibition. *Journal of Experimental Psychology*, **41**, 199-204.

Bjork, R. A., & Richardson-Klavehn, A. (1988). On the puzzling relationship between environmental context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313-344). Hillsdale, NJ: Erlbaum.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, **99**, 229-246.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.

Block, R. A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 530-544.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 85-123). Washington, DC: Winston.

Bower, G. H., & Cohen, P. R. (1982). Emotional influences in memory and thinking: Data and theory. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition* (pp. 263-289). Hillsdale, NJ: Erlbaum.

Cann, A., & Ross, D. A. (1989). Olfactory stimuli as context cues in human memory. *American Journal of Psychology*, **102**, 91-102.

Carr, H. A. (1925). *A study of mental activity*. New York: Longmans, Green.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, **24**, 403-416.

Chappell, M., & Humphreys, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review*, **101**, 103-128.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cousins, R., & Hanley, R. (1996). The effect of environmental context on recall and category clustering scores following relational and individual item processing: A test of the outshining hypothesis. *Memory*, **4**, 79-90.

Dallett, K., & Wilcox, S. G. (1968). Contextual stimuli and proactive inhibition. *Journal of Experimental Psychology*, **78**, 475-480.

Dalton, P. (1993). The role of stimulus familiarity in context-dependent recognition. *Memory & Cognition*, **21**, 223-234.

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

Eckert, E., Kanak, N. J., & Stevens, R. (1984). Memory for fre-

quency as a function of environmental context. *Bulletin of the Psychonomic Society*, **22**, 507-510.
Eich, [J.] E. (1985). Context, memory, and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 764-770.
Eich, [J.] E. (1989). Theoretical issues in state dependent memory. In H. L. Roediger III and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 331-354). Hillsdale, NJ: Erlbaum.
Eich, [J.] E. (1995). Mood as a mediator of place dependent memory. *Journal of Experimental Psychology: General*, **124**, 293-308.
Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.
Farnsworth, P. R. (1934). Examinations in familiar and unfamiliar surroundings. *Journal of Social Psychology*, **5**, 128-129.
Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, **13**, 333-345.
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.
Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, **7**, 95-112.
Glenberg, A. M. (1997). What memory is for. *Behavioral & Brain Sciences*, **20**, 1-55.
Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition*, **26**, 651-658.
Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: Land and underwater. *British Journal of Psychology*, **66**, 325-331.
Godden, D. R., & Baddeley, A. D. (1980). When does context influence recognition memory? *British Journal of Psychology*, **71**, 99-104.
Greenspoon, J., & Ranyard, R. (1957). Stimulus conditions and retroactive inhibition. *Journal of Experimental Psychology*, **53**, 55-59.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.
Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528-551.
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.
Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
Kolers, P. A., & Roediger, H. L., III (1984). Procedures of mind. *Journal of Verbal Learning & Verbal Behavior*, **23**, 425-449.
Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality & Social Psychology*, **49**, 58-69.
Malpass, R. S., & Devine, P. G. (1981). Guided memory in eyewitness identification. *Journal of Applied Psychology*, **66**, 343-350.
Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning & Motivation*, **16**, 398-412.
Mazur, J. E. (1986). *Learning and behavior.* Englewood Cliffs, NJ: Prentice-Hall.
McGeoch, J. A. (1942). *The psychology of human learning*. New York: Longmans.
McSpadden, M., Schooler, J. W., & Loftus, E. F. (1988). Here today, gone tomorrow: The appearance and disappearance of context effects. In D. M. Thomson & G. M. Davies (Eds.), *Memory in context: Context in memory* (pp. 215-230). New York: Wiley.
Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, **104**, 839-862.
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.
Murnane, K., & Phelps, M. P. (1994). When does a different environmental context make a difference in recognition? A global activation model. *Memory & Cognition*, **22**, 584-590.
Murnane, K., & Phelps, M. P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 158-172.
Pan, S. (1926). The influence of contextual conditions upon learning and recall. *Journal of Experimental Psychology*, **9**, 468-491.
Pessin, J. (1932). The effect of similar and dissimilar conditions upon learning and relearning. *Journal of Experimental Psychology*, **15**, 427-435.
Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, **88**, 93-134.
Rand, G., & Wapner, S. (1967). Postural status as a factor in memory. *Journal of Verbal Learning & Verbal Behavior*, **6**, 268-271.
Reed, H. J. (1931). The influence of a change of conditions upon the amount recalled. *Journal of Experimental Psychology*, **14**, 632-649.
Saufley, W. H., Jr., Otaka, S. R., & Bavaresco, J. L. (1985). Context effects: Classroom tests and contextual independence. *Memory & Cognition*, **13**, 522-528.
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. (1984). A comparison of two techniques for reducing context-dependent forgetting. *Memory & Cognition*, **12**, 477-482.
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 D. M. Thomson & G. M. Davies (Eds.), *Memory in context: Context in memory* (pp. 13-34). New York: Wiley.
Smith, S. M. (1994). Theoretical principles of context-dependent memory. In P. Morris & M. Gruneberg (Eds.), *Theoretical aspects of memory* (Aspects of Memory, 2nd ed., Vol. 2, pp. 168-195). New York: Routledge.
Smith, S. M. (1995). Mood is a component of mental context: Comment on Eich (1995). *Journal of Experimental Psychology: General*, **124**, 309-310.
Smith, S. M., Glenberg, A. [M.], & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, **6**, 342-353.
Smith, S. M., Heath, F. R., & Vela, E. (1990). Environmental context-dependent homophone spelling. *American Journal of Psychology*, **103**, 229-242.
Smith, S. M., & Sinha, A. K. (1987). *Effects of brief immersion in a flotation tank on memory and cognition* (Tech. Rep. No. CSCS-004). College Station, TX: Texas A&M University, Committee for the Study of Cognitive Science.
Smith, S. M., & Vela, E. (1986, November). *Outshining: The relative effectiveness of cues*. Paper presented at the 27th Annual Meeting of the Psychonomic Society, New Orleans.
Smith, S. M., & Vela, E. (1992). Environmental context-dependent eyewitness recognition. *Applied Cognitive Psychology*, **6**, 125-139.
Smith, S. M., Vela, E., & Williamson, J. E. (1988). Shallow input processing does not induce environmental context-dependent recognition. *Bulletin of the Psychonomic Society*, **26**, 537-540.

Strand, B. Z. (1970). Change of context and retroactive inhibition. *Journal of Verbal Learning & Verbal Behavior*, **9**, 202-206.

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**, 359-380.

Ucros, C. G. (1989). Mood state-dependent memory: A meta-analysis. *Cognition & Emotion*, **3**, 139-169.

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.

APPENDIX
Studies Included in the Meta-Analysis

Study	Paradigm*	Associative Input Processing	Test Type	Weighted Effect Size (*d*)†	95% Confidence Interval for *d*	
					Lower	Upper
Bell, Hess, Hill, Kukas, Richards, & Sargent (1984)	Reinstatement	Yes	Recall	+.18	−.29	+.65
Bilodeau & Schlosberg (1951)	Interf Reduc	Yes	Cued	+.55	−.03	+1.13
Canas & Nelson (1986)	Reinstatement	Unsure	Recog	+1.06	+.21	+1.92
Cohen, Peterson, & Mantini-Atkinson (1987)	Reinstatement	Unsure	Recall	+.39	−.07	+.86
Cousins & Hanley (1996, Experiment 1)	Reinstatement	Unsure	Recog	+.11	−.33	+.55
Cousins & Hanley (1996, Experiment 2)	Reinstatement	Unsure	Recog	−.27	−.71	+.17
Cutler, Penrod, & Martens (1987)	Imaginal	Unsure	Recog	−.09	−.41	+.23
Dallett & Wilcox (1968, Experiment 1)	Interf Reduc	Unsure	Recall	+1.48	+.58	+2.39
Dallett & Wilcox (1968, Experiment 3)	Interf Reduc	Unsure	Recall	+1.06	+.13	+2.00
Dallett & Wilcox (1968, Experiment 3)	Interf Reduc	Unsure	Recall	+.69	−.13	+1.52
Dallett & Wilcox (1968, Experiment 4)	Interf Reduc	Unsure	Recall	+.74	+.02	+1.45
Dalton (1993, Experiment 1)	Reinstatement	Yes	Recog	+.15	−.38	+.67
Dalton (1993, Experiment 1)	Reinstatement	No	Recog	+1.01	−.71	+.17
Dalton (1993, Experiment 2)	Reinstatement	Yes	Recog	+.17	−.27	+.60
Dalton (1993, Experiment 2)	Reinstatement	No	Recog	+.43	−.01	+.86
DeVane & Parkman (1978)	Reinstatement	Yes	Cued recall	+.08	−.36	+.52
Dolinsky & Zabrucky (1983)	Reinstatement	Unsure	Recall	+.68	−.04	+1.39
Eich (1985)	Reinstatement	No	Recall	+.25	−.24	+.74
Eich (1995, Experiment 2)	Reinstatement	Yes	Recall	+.18	−.39	+.74
Eich (1995, Experiment 3)	Reinstatement	Yes	Recall	−.29	−.86	+.28
Emmerson (1986)	Reinstatement	Unsure	Recog	+.66	−.05	+1.37
Fernandez & Glenberg (1985, Experiment 1)	Reinstatement	Yes	Recall	+.12	−.45	+.68
Fernandez & Glenberg (1985, Experiment 2)	Reinstatement	Yes	Recall	+.15	−.65	+.95
Fernandez & Glenberg (1985, Experiment 3)	Reinstatement	Yes	Recall	−.31	−.80	+.19
Fernandez & Glenberg (1985, Experiment 4)	Reinstatement	Yes	Recall	−.19	−.68	+.30
Fernandez & Glenberg (1985, Experiment 5)	Reinstatement	Yes	Recall	+.34	−.36	+1.04
Fernandez & Glenberg (1985, Experiment 6)	Reinstatement	Yes	Recall	−.46	−1.17	+.24
Fernandez & Glenberg (1985, Experiment 7)	Reinstatement	Yes	Recall	−.09	−.65	+.48
Fernandez & Glenberg (1985, Experiment 8)	Reinstatement	Unsure	Recall	+.26	−.31	+.83
Fisher, Geiselman, MacKinnon, & Holland (1984)	Imaginal	Unsure	Recall	+.27	−.24	+.77
Glenberg (1979, Experiment 3)	Mult Context	Unsure	Recall	+.84	+.12	+1.57
Glenberg (1979, Experiment 3)	Reinstatement	Unsure	Recall	+.19	−.51	+.88
Godden & Baddeley (1975)	Reinstatement	Unsure	Recall	+.91	−.06	+1.88
Godden & Baddeley (1980)	Reinstatement	Unsure	Recog	0.00	−1.13	+1.13
Greenspoon & Ranyard (1957)	Interf Reduc	Unsure	Recall	+1.75	+1.18	+2.32
Humphreys, Pike, Bain, & Tehan (1988)	Reinstatement	Unsure	Recog	−.16	−.72	+.41

APPENDIX (Continued)

Study	Paradigm*	Associative Input Processing	Test Type	Weighted Effect Size (d)†	95% Confidence Interval for d	
					Lower	Upper
Jacoby (1983)	Reinstatement	No	Recog	0.00	−.80	+.80
Jensen, Dibble, & Anderson (1971)	Interf Reduc	Yes	Recall	+.15	−.47	+.77
Jensen, Harris, & Anderson (1971)	Reinstatement	No	Recall	+. 28	+.02	+.54
Krafka & Penrod (1985)	Imaginal	Unsure	Recog	+1.71	+1.21	+2.21
Malpass & Devine (1981)	Imaginal	Unsure	Recog	+.49	+.03	+.96
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 1)	Reinstatement	Yes	Recall	+.21	−.28	+.70
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 2)	Reinstatement	Unsure	Recall	+.61	−.04	+1.19
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 3)	Reinstatement	Yes	Recall	+.46	−.02	+.90
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 4)	Reinstatement	Unsure	Recall	−.07	−.50	+.36
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 4)	Reinstatement	Yes	Recall	−.02	−.45	+.41
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 5)	Reinstatement	Unsure	Recall	+.49	−.22	+1.19
McDaniel, Anderson, Einstein, & O'Halloran (1988, Experiment 5)	Reinstatement	Yes	Recall	−.80	−1.52	−.08
McSpadden, Schooler, & Loftus (1988, Experiment 1)	Imaginal	Unsure	Cued recall	+.35	+.07	+.64
McSpadden, Schooler, & Loftus (1988, Experiment 1)	Imaginal	Unsure	Cued recall	−.21	−.68	+.26
McSpadden, Schooler, & Loftus (1988, Experiment 3)	Imaginal	Unsure	Cued recall	+.33	−.13	+.80
Nagge (1935, Experiment 2)	Interf Reduc	Yes	Serial recall	+.07	−.49	+.64
Nixon & Kanak (1981)	Reinstatement	Unsure	Recall	+.07	−.42	+.56
Sanders (1984)	Reinstatement	Unsure	Recog	−.16	−.85	+.53
Smith (1979, Experiment 1)	Reinstatement	Unsure	Recall	+.91	+.11	+1.70
Smith (1979, Experiment 2)	Reinstatement	Unsure	Recall	+1.01	+.08	+1.94
Smith (1979, Experiment 3)	Imaginal	Unsure	Recall	+.88	−.04	+1.79
Smith (1982, Experiment 1)	Mult Context	Unsure	Recall	+.68	+.18	+1.18
Smith (1982, Experiment 2)	Mult Context	Unsure	Recog	−.03	−.72	+.66
Smith (1982, Experiment 3)	Mult Context	Unsure	Recall	+.32	−.18	+.83
Smith (1982, Experiment 3)	Reinstatement	Unsure	Recall	+.64	+.13	+1.16
Smith (1984)	Mult Context	Unsure	Recall	+.19	−.32	+.70
Smith (1984)	Imaginal/MC	Unsure	Recall	+.27	−.24	+.77
Smith (1985a, Experiment 1)	Reinstatement	Unsure	Recog	+.66	+.02	+1.29
Smith (1985a, Experiment 2)	Reinstatement	Unsure	Recall	+.18	−.26	+.62
Smith (1985a, Experiment 2)	Reinstatement	Unsure	Recog	+.48	+.03	+.92
Smith (1985a, Experiment 2)	Reinstatement	Unsure	Cued Rec	+.26	−.18	+.70
Smith (1985a)	Mult Context	Yes	Recall	+.45	+.12	+.78
Smith (1985b, Experiment 1)	Reinstatement	Unsure	Recall	+.81	+.23	+1.40
Smith (1985b, Experiment 2)	Reinstatement	Unsure	Recall	+1.11	+.56	+1.65
Smith (1986, Experiment 1)	Reinstatement	No	Recall	+.87	+.18	+1.55
Smith (1986, Experiment 2)	Reinstatement	No	Recog	+.38	+.02	+.75
Smith (1986, Experiment 3)	Reinstatement	No	Recog	+.39	−.25	+1.04
Smith (1986, Experiment 3)	Reinstatement	Yes	Recog	−.05	−.69	+.59
Smith, Glenberg, & Bjork (1978, Experiment 1)	Mult Context	Unsure	Recall	+.86	+.13	+1.58
Smith, Glenberg, & Bjork (1978, Experiment 2)	Interf Reduc	Yes	Recall	+.75	+.04	+1.47
Smith, Glenberg, & Bjork (1978, Experiment 3)	Reinstatement	Yes	Recall	+1.02	+.08	+1.95
Smith, Glenberg, & Bjork (1978, Experiment 4)	Reinstatement	Unsure	Recall	+1.23	+.52	+1.94
Smith, Glenberg, & Bjork (1978, Experiment 4)	Reinstatement	Unsure	Recog	+.14	−.30	+.58

APPENDIX (Continued)

Study	Paradigm*	Associative Input Processing	Test Type	Weighted Effect Size (d)†	95% Confidence Interval for d Lower	Upper
Smith, Glenberg, & Bjork (1978, Experiment 5)	Reinstatement	Unsure	Recog	−.11	−.60	+.38
Smith & Vela (1992)	Reinstatement	Unsure	Recog	+.84	+.49	+1.18
Smith & Vela (1992)	Imaginal	Unsure	Recog	−.12	−.45	+.21
Smith, Vela, & Williamson (1988, Experiment 1)	Reinstatement	No	Recog	+.13	−.28	+.54
Smith, Vela, & Williamson (1988, Experiment 1)	Reinstatement	Yes	Recog	0.00	−.41	+.41
Smith, Vela, & Williamson (1988, Experiment 2)	Reinstatement	No	Recog	+.18	−.24	+.60
Smith, Vela, & Williamson (1988, Experiment 2)	Reinstatement	Yes	Recog	+.45	+.02	+.88
Smith, Vela, & Williamson (1988, Experiment 3)	Reinstatement	No	Recog	+.59	−.04	+1.23
Smith, Vela, & Williamson (1988, Experiment 4)	Reinstatement	No	Recog	+.12	−.30	+.55
Smith, Vela, & Williamson (1988, Experiment 4)	Reinstatement	Yes	Recog	+.25	−.18	+.67
Strand (1970)	Interf Reduc	Yes	Recall	+.07	−.59	+.72
Wilhite (1991, Experiment 1)	Reinstatement	Yes	Recall	−.30	−.85	+.26
Wilhite (1991, Experiment 1)	Reinstatement	Unsure	Recall	−.72	−1.23	−.22
Wilhite (1991, Experiment 2)	Reinstatement	Unsure	Recall	−.51	−1.02	−.01

*Studies coded as "Imaginal" were analyzed separately. †Positive effect sizes reflect mean differences which favor a) context reinstatement for first-order paradigms, b) interference reduction as a consequence of context variation, or c) increased memory performance for a single list or multiple lists as a function of multiple context encoding. Each effect size is weighted by the inverse of its variance (Hedges & Olkin, 1985).

REFERENCES FOR THE META-ANALYSIS

Bell, P. A., Hess, S., Hill, E., Kukas, S. L., Richards, R. W., & Sargent, D. (1984). Noise and context-dependent memory. *Bulletin of the Psychonomic Society*, **22**, 99-100.

Bilodeau, I. M., & Schlosberg, H. (1951). Similarity in stimulating conditions as a variable in retroactive inhibition. *Journal of Experimental Psychology*, **41**, 199-204.

Canas, J. J., & Nelson, D. L. (1986). Recognition and environmental context: The effect of testing by phone. *Bulletin of the Psychonomic Society*, **24**, 407-409.

Cohen, R. L., Peterson, M., Mantini-Atkinson, T. (1987). Inter-event differences in event memory: Why are some events more recallable than others? *Memory & Cognition*, **15**, 109-118.

Cousins, R., & Hanley, R. (1996). The effect of environmental context on recall and category clustering scores following relational and individual item processing: A test of the outshining hypothesis. *Memory*, **4**, 79-90.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology*, **72**, 629-637.

Dallett, K., & Wilcox, S. G. (1968). Contextual stimuli and proactive inhibition. *Journal of Experimental Psychology*, **78**, 475-480.

Dalton, P. (1993). The role of stimulus familiarity in context-dependent recognition. *Memory & Cognition*, **21**, 223-234.

Devane, J. R., & Parkman, R. A. (1978). Effects of imagery instructions when the context is changed. *Perceptual & Motor Skills*, **46**, 716-718.

Dolinsky, R., & Zabrucky, K. (1983). Effects of environmental context changes on memory. *Bulletin of the Psychonomic Society*, **21**, 423-426.

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

Eich, [J.] E. (1995). Mood as a mediator of place dependent memory. *Journal of Experimental Psychology: General*, **124**, 293-308.

Emmerson, P. G. (1986). Effects of environmental context on recognition 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.

Fisher, R. P., Geiselman, R. E., MacKinnon, D. P., & Holland, H. L. (1984). Hypnotic and cognitive interviews to enhance the memory of eyewitnesses to crime. *International Journal of Investigative & Forensic Hypnosis*, **7**, 28-31.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, **7**, 95-112.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: Land and underwater. *British Journal of Psychology*, **66**, 325-331.

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

Greenspoon, J., & Ranyard, R. (1957). Stimulus conditions and retroactive inhibition. *Journal of Experimental Psychology*, **53**, 55-59.

APPENDIX (Continued)

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 an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.
Jensen, L. [C.], Dibble, J., & Anderson, D. C. (1971). Effects of a contextual change upon retroactive inhibition. *Psychological Reports*, **29**, 39-46.
Jensen, L. C., Harris, K., & Anderson, D. C. (1971). Retention following a change in ambient contextual stimuli for six age groups. *Developmental Psychology*, **4**, 394-399.
Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality & Social Psychology*, **49**, 58-69.
Malpass, R. S., & Devine, P. G. (1981). Guided memory in eyewitness identification. *Journal of Applied Psychology*, **66**, 343-350.
McDaniel, M. A., & Anderson, C. D., Einstein, G. O., & O'Halloran, C. M. (1988). Modulation of environmental reinstatement effects through encoding strategies. *American Journal of Psychology*, **102**, 523-548.
McSpadden, M., Schooler, J. W., & Loftus, E. F. (1988). Here today, gone tomorrow: The appearance and disappearance of context effects. In D. M. Thomson & G. M. Davies (Eds.), *Memory in context: Context in memory* (pp. 215-230). New York: Wiley.
Nagge, J. W. (1935). An experimental test of the theory of associative interference. *Journal of Experimental Psychology*, **18**, 663-682.
Nixon, S. J., & Kanak, N. J. (1981). The interactive effects of instructional set and environmental context changes on the serial position effect. *Bulletin of the Psychonomic Society*, **18**, 237-240.
Sanders, G. L. (1984). Effects of context cues on eyewitness identification responses. *Journal of Applied Social Psychology*, **14**, 386-397.
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. (1984). A comparison of two techniques for reducing context-dependent forgetting. *Memory & Cognition*, **12**, 477-482.
Smith, S. M. (1985a). Effects of number of study environments and learning instructions on free-recall clustering and accuracy. *Bulletin of the Psychonomic Society*, **23**, 440-442.
Smith, S. M. (1985b). 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., & Glenberg, A. [M.], & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, **6**, 342-353.
Smith, S. M., & Vela, E. (1992). Environmental context-dependent eyewitness recognition. *Applied Cognitive Psychology*, **6**, 125-139.
Smith, S. M., Vela, E., & Williamson, J. E. (1988). Shallow input processing does not induce environmental context-dependent recognition. *Bulletin of the Psychonomic Society*, **26**, 537-540.
Strand, B. Z. (1970). Change of context and retroactive inhibition. *Journal of Verbal Learning & Verbal Behavior*, **9**, 202-206.
Wilhite, S. C. (1991). Evidence of a negative environmental reinstatement effect. *British Journal of Psychology*, **82**, 325-342.

(Manuscript received August 12, 1999;
revision accepted for publication September 8, 2000.)