GLOBAL VERSUS LOCAL: FUNCTIONAL COMPONENTS OF SCENE CONTEXT EFFECTS IN RECALL

By

KIMELE PERSAUD

A thesis submitted to the

Graduate School- New Brunswick

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Masters of Science

Graduate Program in Psychology

Written under the direction of

Dr. Pernille Hemmer

And approved by

____

____

____

New Brunswick, New Jersey

January, 2015

ABSTRACT OF THE THESIS

Global versus Local: Functional Components of Scene Context Effects in Recall

by KIMELE PERSAUD

Thesis Director:

Dr. Pernille Hemmer

Expectation for context is perhaps the most influential contributor to episodic memory. Although research has investigated the influence of functional components of scene context in perception, little is known about the independent contributions of these components to long-term episodic memory. In this investigation we find that different from perception, these components make substantially different contributions to memory. Namely, the global context component that binds objects to natural scenes is important at short study times, where a lack of global binding information appears to disrupt meaning extraction. Local context components that bind objects to each other within a scene (i.e., spatial and associative) are important for sustaining memory performance. Disrupted spatial information forces a longer visual search, which requires more study time for effective encoding. Lacking associative information has a detrimental effect on recall following short study times—effectively equating performance to short term memory. This has important theoretical implications.

Keywords: Context; Episodic Memory; Long-Term Memory; Prior Knowledge; Natural Scenes; Objects in Scenes.

TABLE OF CONTENTS

1. Introduction

Expectations play an important role in how we think, reason, and interact with the world around us. Expectations based on the regularities of our environment can help us perform a variety of tasks, such as perceiving and categorizing novel objects, recalling events from memory, and even making predictions about the future (Brady & Oliva, 2008; Griffiths & Tenenbaum, 2006; Hemmer & Persaud, 2014; Hemmer, Steyvers, & Miller, 2010; Hemmer, Tauber, & Steyvers, in revision; Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Vevea 2000; Jern & Kemp, 2013; and Persaud & Hemmer, 2014). These expectations become central when we are faced with everyday tasks, such as recalling objects in a room, where we can use expectations learned over time. For example, imagine you were asked to tell your friend about the hotel room on your last vacation, which may be complicated by the amount of time that has passed. Given that you have expectations for the contextual relationships between hotel rooms and objects—e.g. hotels have beds, bedside tables, and alarm clocks—you can use this information to help make relatively accurate guesses about the objects in the hotel room you are trying to recall. Expectations for contextual relationships between objects and scenes often include other pertinent information such as the associative relationship between objects (e.g., if there is a bed, there is likely a nightstand), or the proximity with which certain objects are located relative to one another (e.g. alarm clocks sit on nightstands). What remains unclear is how these different components of expectations influence our ability to remember long-term.

Expectations for scenes have important implications for how we visually search, allocate attention, and recognize objects in scenes, (Bar, 2004; Bar & Ullman, 1996;

Inoue & Takeda, 2012; Mandler & Johnson, 1976; Palmer, 1975). Expectations support quick scene interpretation, even for severely impoverished images (Biederman et. al., 1982; Torralba, 2003), and help make object recognition more accurate (Galleguillos & Belongie, 2010). In addition, people use these expectations to aid recall (e.g., Hemmer & Steyvers, 2009b, for a review see Hemmer & Persaud, 2014).

The benefit of having expectations for scene context results from knowledge of the functional components that form scene context, and are defined by the global and local structure (Galleguillos & Belongie, 2010; Torralba, 2003). The global component refers to the overall configuration of a scene. Global information, or global context, is responsible for the unification of objects and the background (see Figure 1a for scenes with global context preserved and partially removed) and supports quick semantic interpretation of a scene (Potter, Staub, Rado, & O'Connor, 2002; Torralba, 2003). It also affords individuals the ability to make predictions about objects that prototypically accompany the scene type (Galleguillos & Belongie, 2010).

Similarly, the local component refers to the relations among objects in scenes, or particular regions of scenes. Local information, or local context, is derived from the associative and spatial arrangements of scene objects. These arrangements enhance scene perception and object recognition (Biederman, 1972; Biederman et al., 1982; Galleguillos & Belongie, 2010; Palmer, 1975; Torralba, 2003). Take, for example, the ambiguous object in Figure 1b (left panel). The intrinsic properties of this object make it hard to recognize. However, placed next to another object with which it is often associated (i.e., the trashcan), it becomes clear that the object is a crumpled paper.

The spatial relationships of objects in scenes also contribute to object recognition. Figure 1c (left panel) demonstrates an ambiguous object placed above a table, which leads one to conclude that it is a tablecloth. However, that same object placed below the table (Figure 1c, middle panel), leads one to conclude that it is an area rug. When the object is placed above a bed (Figure 1c, right panel), it appears to be a blanket. Taken together, scene context is comprised of global context, which unifies the objects and background, and local context, which determines the associative and spatial relationships among objects[1].

Although it is clear that global and local components of context facilitate scene perception, object identification, and categorization, what remains unclear is how they influence long-term episodic memory for scenes. Previous research investigating the influence of scene context on memory has either employed short-term or working memory to understand scene perception (Biederman, 1972; Hollingworth & Henderson, 1998), or did not distinguish the influence of global and local context (Hemmer & Steyvers, 2009; Steyvers & Hemmer, 2012). Here, we present an investigation of the relative contribution of global and local components of context on long-term episodic memory. We extend the work of Hemmer and Steyvers (2009), which assessed the contribution of prior expectations for context on episodic memory for natural scenes. In a series of studies, Hemmer and Steyvers quantified peoples' expectations and memory for objects in full context natural scenes (Figure 1a, left panel).

---

[1] Some information referred to as local context—e.g., associative relationships—may also be captured in the global structure. However, the goal here is to evaluate them as separate contributions.

Prior expectations for objects in scenes were assessed by asking participants to list all objects that make up five natural scene types (i.e., kitchen, dining, hotel, urban, and office). Interestingly, they found that by simply guessing with context-based expectations, participants achieved a high degree of accuracy when responses were scored against actual images. Free recall for the same natural scenes was then tested as a function of study time. The results showed high accuracy for both short and long study

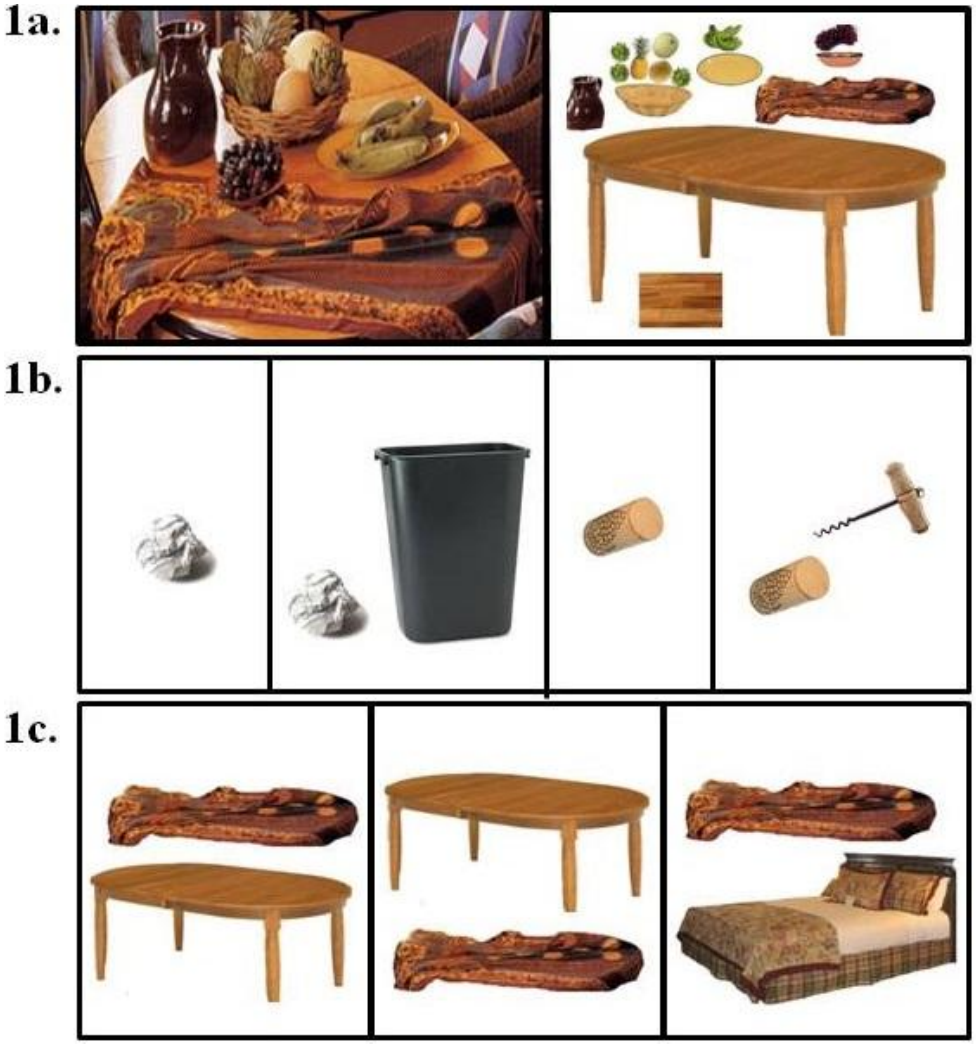


Figure 1. 1a shows a natural scene – a dining room – with the global context preserved (left) and removed (right). 1b shows the associative relationship among objects. 1c shows the spatial relationship among objects.

times. Here, we use novel simulated natural scenes based on Hemmer and Steyvers, in which we systematically remove global and local context to measure their influences on memory. We also sought to investigate whether other factors (e.g. additional study time) could compensate for removed contextual features.

2. Experiments

Figure 2. Left: full natural dining room scene. Right: partial-context dining room scene, in which the background was removed.

In three experiments, we systematically evaluated the effect of the functional components of context on episodic memory. The effect of global context was evaluated in Experiment 1, and the effects of local context in Experiments 2 and 3. Performance in all experiments was evaluated against the prior expectation experiment in Hemmer and Steyvers (2009)—hereafter referred to as the Original Study, where, responses were based only on contextual knowledge and expectations, and not episodic memory. We predicted that accuracy would decrease as a function of the incremental exclusion of global and local context. The same basic components of stimuli were used across all experiments—namely individual objects from natural scenes presented on a white background—with manipulations varied between experiments. We first outline the general methods for generating the stimuli, correcting free recall responses, and analyzing responses for all experiments. We then present the methods and results for each experiment.

2.1 General Methods

2.1.1 Materials

To create the stimuli for the experiments, we used the original images, along with a ground truth assessment for objects truly present in the scenes from the Original Study. Every object named in the ground truth was then cropped from the original image using Paint and Photoshop (Figure 2, right panel). If an object could not be clearly cropped, Google Images was used to find an object that closely matched the original scene. A panel of three students rated the Google items compared to the original items (across all experimental conditions rater agreement was 97%). Individual objects were then placed on a white background in an order that differed for each experiment as outlined in each Materials and Procedure section.

Extensive measures were taken to ensure that the context manipulated scenes matched the original scene, which included: making the most salient objects in the original scene the most salient object in the context-disrupted scene, matching up the sizes of the objects, making the objects as clear as possible, and matching up the colors and angles of the objects as closely as possible. The size and saliency of the objects was especially important when creating the context manipulated scenes because people are known to have a "normative viewing size" preference for a given object (Konkle & Oliva, 2007).

2.1.2 Response Normalization

Recall responses for all experiments were corrected for spelling, plurals, capitalization, and qualifiers (e.g., numbers, and color). For example, "chair" and "chairs" were mapped to the single entry "chair", and "silver car" was mapped to "car". All short form responses were corrected to the full word (e.g., mayo was mapped to mayonnaise, and fridge was mapped to refrigerator). Responses that could not be interpreted were removed. The correction rules were automated and applied to all datasets uniformly.

2.1.3. Ground Truth

To measure performance in all experiments, we checked whether a recalled object was part of any of the responses given in the ground truth assessment of the Original Study, measured by asking subjects to "report as many objects as you can see" for each image. If a response given in our memory studies matched the ground truth, it was scored as correct. If not, we manually checked whether the recalled object could be considered as a description of an object in the image. This object was then added to the original ground truth list. Only if the response still did not match was it scored as incorrect.

2.2 Experiment 1: Partial scene context

We sought to investigate the effect of global context—in the form of scene background—on episodic memory. That is, to what extent does the natural setting of a room (e.g., the walls and ceiling) contribute to successful memory? We predicted that the absence of global context would disrupt quick scene interpretation and would decrement performance, especially at shorter study times. We expected that additional study time would improve memory performance, and would restore accuracy to levels achieved in memory for the non-disrupted scenes. We refer to this as the partial-context condition (Figure 2, right panel).

2.2.1Participants

Fifty-three undergraduate students at Rutgers University participated in exchange for course credit or monetary compensation of $10. A sample size of 45-55 participants was based on the number of participants in the Original Study (Hemmer & Steyvers, 2009) with a minimum of 50% more participants required for the additional study time manipulation (15 seconds), in order to reach equivalent number of participants per image. The same sample size applied to all experiments.

2.2.2. Materials and Procedure

To create the partial-context images, objects were placed onto a white background in the same spatial organization as the objects in the original image. The 10 original images (two for each of five scene types—kitchen, dining room, office, hotel room, and urban) were used to form two sets of five study images. We followed the exact experimental procedure of the Original Study, and employed a recall paradigm in which images were presented at the center of the computer screen for either 2, 10, or 15 seconds. A simple 'find 5 mistakes' picture distracter task was inserted between study and test trials. At test, participants were asked to type all the objects they could recall from the image presented in the preceding study trial. Participants were given clear verbal instructions to ensure that they understood the task.

On study trials, study times were randomly assigned as either 2, 10 or 15 seconds following a Latin square design. On test trials, participants were required to type responses or wait 60 seconds before they could move to the next study trial. Each participant only saw 5 images, one from each scene type, to avoid carryover effects where the memory from one scene type affects recall of another image of the same type. The 5 images were presented in random order. At the end of each of the five test blocks, participants received feedback on the number of correct responses, and how many more objects they could have recalled, in order to encourage participants to stay engaged in the task.

2.2.3. Results

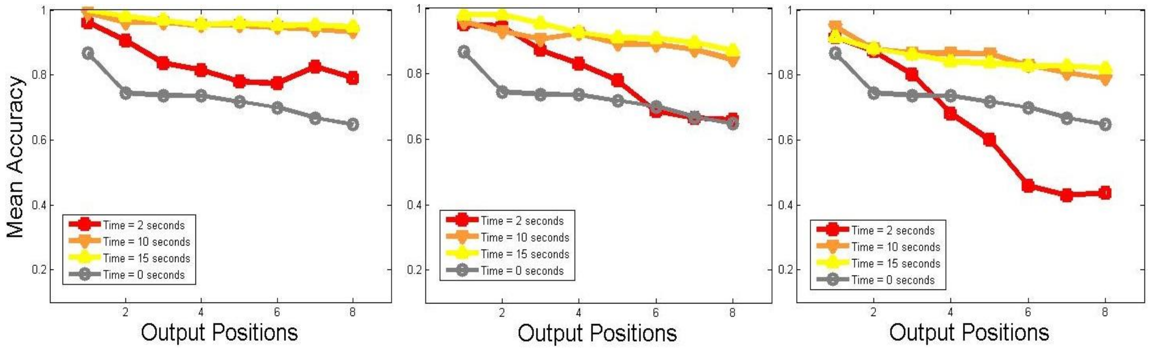


Figure 3. Mean accuracy as a function of output position and study time. Each line gives a different study time condition. The gray line gives performance from the Prior Knowledge condition of Hemmer and Steyvers (2009). Left panel: partial-context condition. Middle panel: no-spatial-context condition. Right panel: random-context condition.

Performance was measured in terms of mean accuracy as a function of the output position and study time (See Figure 3, left panel). Figure 3 also includes the results from the prior expectation condition in the Original Study (gray line), as a baseline for comparison. A one-way ANOVA was conducted to evaluate the effect of study time on recall accuracy averaged across the first eight output positions. Because participants were free to terminate responding at any time, the number of responses differed both between subjects and between study time conditions. We include the first eight responses in the analysis because this allowed for the inclusion of the most responses while also reducing the variability across subjects and conditions, and for consistency across experiments. For further analysis of the number of responses see the section “Number of responses by condition”. There was as a significant main effect of study time (*F* [2, 158] = 11.42, *p*<.001), such that greater accuracy was achieved for longer study times compared with shorter times. Planned post hoc comparisons revealed a significant difference in accuracy between 2 and 10 seconds of study time (*F* [2, 105] = 11.18, *p*<.001), 2 and 15 seconds (*F* [2, 105] = 13.78, *p*<.001), but not between 10 and 15 seconds (*F* [2, 105] = .45, *p*>.05).

Performance in the 2 second condition was significantly reduced relative to both the 10 and 15 second conditions. This suggests that lack of global context disrupts gist extraction that allows for quick scene interpretation and is known to be a fast, early process in scene perception. The removal of global context that binds the object in the scene might slow down the ability to quickly interpret the scene and therefore hurt memory performance at shorter study times. However, this also suggests that 10 seconds is ample study time to recover from the detriment to memory caused by the removal of background context.

2.3 Experiment 2: No-spatial-context

In the next experiment, we investigate the effect of local context—in the form of spatial context—on episodic memory. That is, to what extent does the spatial relationship among objects in a natural context contribute to successful memory? Based on the finding that the removal of the global context (i.e. background wall and ceilings) reduced accuracy at short study times, we predicted that the further absence of local spatial context would decrease accuracy across all study times. We refer to this as the no-spatial-context condition.

2.3.1 Participants

Fifty Rutgers University students participated in exchange for either monetary compensation of $10 or course credit, and were not involved in Experiment 1.

2.3.2 Materials and Procedure

The materials were identical to those used in Experiment 1, except the spatial relationship among objects was not retained. Instead objects within each scene were placed onto the white background in random order. Figure 4 (left panel) shows a sample image from the no-spatial context-condition. The procedure was identical to Experiment 1.

2.3.3 Results

As in Experiment 1, performance was measured in terms of mean accuracy as a function of study time. Figure 3 (middle panel) shows accuracy across study times, as well as the prior knowledge condition for comparison. A one-way ANOVA was conducted to evaluate the effect of study time on recall accuracy averaged across the first eight output positions. There was a significant main effect of study time (*F* [2, 148] = 3.16, p<.05), such that greater accuracy was achieved for longer study times compared with shorter times. Planned post hoc comparisons revealed a significant difference in accuracy only between 2 and 15 seconds of study time (*F* [1, 99] = 6.77, *p*<.01), while there were no significant differences between 2 and 10 seconds or 10 and 15 seconds.

Figure 4. Left panel: No-spatial-context dining room scene in which the spatial relationship between the objects was disrupted. Right panel: Random-context scene, in which objects were randomly selected from the 5 images in each scene set to create the random study list of images.

However, a closer inspection of the graph, particularly at later output positions, suggested a clear trend in the data in the 2 second study condition. Further analysis after the third output position (i.e. positions 4-8) revealed a significant difference in performance in the 2 and 10 second study times (*F* [1, 83] = 13.06, *p*<.001). The graph also suggests that performance in the 2 second condition might not be greater than that of the prior knowledge condition (Hemmer & Steyvers, 2009) illustrated by the gray line. A t-test between the prior knowledge condition and the 2 second study condition averaged over the first eight output positions, however, revealed a significant difference between the 2 second and prior knowledge condition (gray line) averaged across eight output positions (*t* [57] = -3.30, *p* <.001), such that performance was better in the 2 second study condition. A second t-test restricted to the latter output positions (positions 6-8;), however, found no significant difference between the prior knowledge condition and the 2 second condition (*t* [21] = -0.26, *p* =0.798).

These findings suggest that when the spatial relationships among objects in a scene are disrupted, for initial output positions at all study times, people can still employ what equates to short term memory of 3 or 4 items. However, after the third output position, performance at shorter study time quickly drops off and becomes significantly different from memory performance at longer study times. In fact, at later output positions for shorter study times (positions 6-8), there is no difference in performance from guessing with prior knowledge (gray line). This also suggests that at longer study times, people might still search the scene and deduce associative information to sustain memory performance. This is consistent with research showing that a lack of proper spatial relations among the objects in a scene increases response times, as well as error rates in the recognition of individual objects (Bar & Ullman, 2004).

2.4 Experiment 3: Random scene context

The preceding studies revealed a continuous decline in accuracy with the progressive removal of global and local spatial information. In the next experiment, we sought to further investigate the effect of local context—in the form of associative information— on episodic memory, and to quantify pure episodic memory as a basis for comparison. That is, to what extent does the associative relationship among object in a scene contribute to successful memory? To measure this contribution, we tested memory for a scene of random objects in which both global and local information were removed (Figure 4, right panel). We predicted that this absence of natural context would further decrement performance across study times. We refer to this as the random-context condition.

2.4.1 Participants

Forty-eight undergraduate students from Rutgers University participated in exchange for either monetary compensation of $10 or course credit. These participants were not involved in Experiments 1 or 2.

2.4.2 Materials and Procedure

The materials were identical to those used in Experiments 1 and 2, except the scene context among objects was not retained. Instead, objects within each study set were drawn at random from across the 5 scene types, and placed in random order on the white background. In this way, the stimulus no longer retained the global or local context of a natural scene. Figure 4 (right panel) shows a sample image from the random-context condition.

Objects were matched for size from the stimuli in Experiments 1 and 2, such that small and large objects occurred in all scenes, and no one object was allowed to repeat across the 5 random scenes in each set. Again, a ratings panel of three students was used to determine the consistency of the overall quality of the ‘scene’ relative to the previous experimental stimuli. The procedure was identical to Experiments 1 and 2.

2.4.3 Results

As in Experiments 1 and 2, performance was measured in terms of mean accuracy as a function of output position. Figure 3 (right panel) shows accuracy across output position and study time, as well as the results from the prior knowledge condition as a baseline for comparison. A one-way ANOVA was conducted to assess the effect of study time on recall accuracy averaged across the first eight output positions. There were no significant differences borne out in the recall accuracy for the random condition and planned post hoc comparisons revealed only a marginally significant difference between the 2 and 10 second conditions (F [1, 95] = 2.97, p=.08). However, a closer inspection of the graph, particularly at later output positions, suggested a clear trend in the data, particularly in the 2 second condition. Therefore, we looked at the number of responses given by subjects in this condition and found that 5 subjects only provided 2 responses and 1 subject only provided 1 response. Data from these subjects comprised only 11 data points, which were excluded from the subsequent analysis. A one-way ANOVA, excluding these subjects, revealed a significant main effect of study time (*F* [2, 122] = 3.83, p<.05), where greater accuracy was achieved for longer study times. Post hoc comparisons showed a significant difference in mean recall accuracy between the 2 and 10 second (*F* [1, 81] = 6.18, $p<.01$), and 2 and 15 second (*F* [1, 81] = 3.93, $p<.05$), but not between the 10 and 15 second study time conditions. The graph also shows a relationship between the 2 second condition and prior knowledge (gray line) such that after the 4th output position, performance in the 2 second condition is far below that of the prior knowledge condition. (Since prior knowledge, illustrated in this line, is not equivalent to prior knowledge for the stimuli in this experiment where all global and local information is removed, we will not provide a formal statistical comparison.)

With the removal of associative relationships among objects in a scene, accuracy for all study times decreased. Performance in the 2 second condition was significantly impaired relative to both the 10 and 15 second conditions—and even relative to the prior knowledge condition—particularly at later outputs. The removal of both global and local context disrupts the binding of objects not only to the scene, but to each other, which is more taxing on the memory system. However, the lack of difference between the 10 and 15 second conditions suggests that while 10 seconds is enough to improve performance above the threshold of prior knowledge, the additional 5 seconds provided in the 15 second condition is not enough to recover to the levels of performance in the previous experiments where contextual information was available. A detailed analysis of performance across experiments will be provided in the next section.

2.5 General Results

2.5.1 Comparing accuracy across context conditions

The results from the three experiments appear to show a proportional decline of memory accuracy with the successive removal of global and local scene context. To evaluate this effect, a repeated measures 2-way ANOVA was conducted to compare performance for each study time (i.e., 2, 10, and 15 seconds) across the three experiments. There was a significant main effect of experimental condition. (*F* [2, 431] = 13, *p*=.001), and a significant main effect of study time (*F* [2, 431] = 15.06, *p*<.001). There was no significant interaction. Planned post hoc comparisons of study time between experiments revealed the following pattern of performance: there was no significant difference in performance between the partial and no-spatial context conditions (*t* [101] = -0.49, *p* =0.62), but there was a significant decrease between the no-spatial and random conditions (*t* [89] = 2.77, *p* <.01). In the 10 second condition, there was a significant decrease in performance between the partial and no-spatial conditions (*t* [101] = 2.61, *p* <.01), but not between the no-spatial and random conditions (*t* [89] = 0.70, *p* =0.49). Lastly, in the 15 second condition there was a significant decrease in performance between partial and no-spatial-context (*t* [101] = 2.12, *p* <.01), and between no-spatial and random-context (*t* [89] = 3.51, *p* <.001).

The removal of global context (Experiment 1) hurts performance at shorter study times, but can be compensated for by longer study times. Performance at shorter study times is further disrupted with the removal of associative information (Experiment 3), suggesting that context information that binds objects to scenes and objects to each other most affects performance at shorter study times. Performance in both the 10 and 15 second conditions was negatively affected by the removal of spatial information (Experiment 2). This might be due, in part, to the fact that performance in these two study conditions was near ceiling in the partial-context condition.

2.5.2 Number of responses across experimental conditions

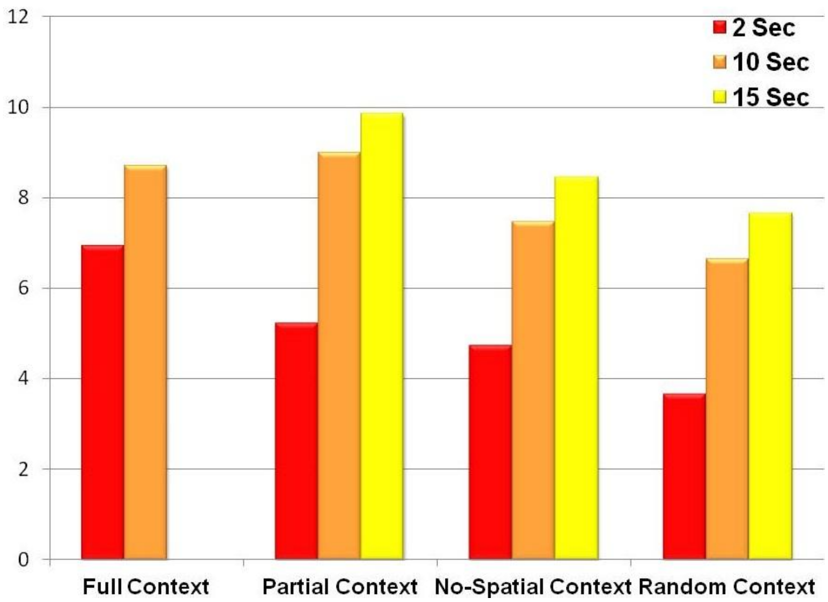


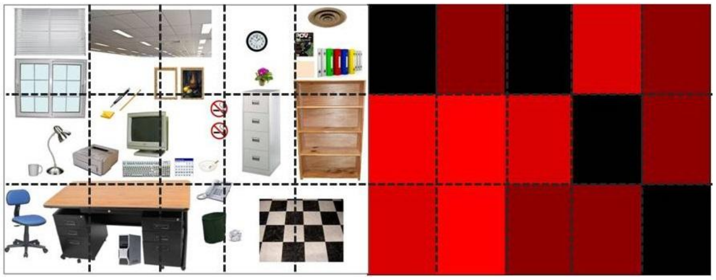
Figure 5. Average number of responses by context condition and study time.

The availability of global and local context also affected the number of items recalled. Figure 5 shows the number of responses as a function of context condition and study time. A 2-way repeated measures ANOVA was conducted to evaluate the effect of study time (2, 10, and 15 seconds) and context condition (Experiment 1, 2, and 3) on number of responses. There was a significant main effect of study time (*F* [2, 294] = 210, *p*<.001), with more responses for longer study times. There was a significant main effect of context condition (*F* [2, 294] = 45.39, *p*<.001), with more responses for more available context. Lastly, there was a significant interaction between study time and context condition (*F* [4, 294] = 6.67, *p*<.001), such that there were more responses given in the 10 and 15 second conditions compared to the 2 second condition as a function of the context manipulations.

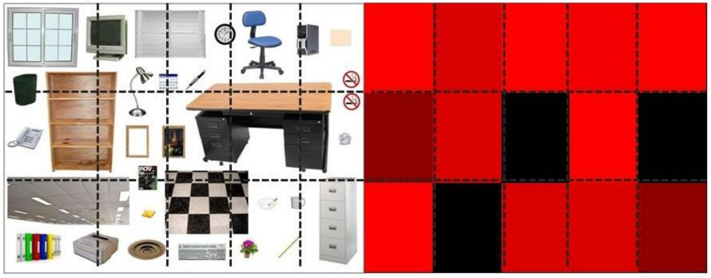
2.5.3 Evaluating response regions within scenes across experimental conditions

The removal of global context appears to inhibit the ability to quickly extract the meaning of a scene. The removal of local spatial information further reduced recall performance. This might be due to the need to search the scene to ensure that objects are present in a scene. In full context natural scenes, objects are often located in close proximity based on use (e.g., in an office, a, pen and notepad might appear together on a desk). However, once the local spatial information is removed, this is no longer the case, and one cannot assume that just because there is a pen in a scene there will also be a notepad and a pencil. This information can only be obtained by searching the scene. To assess recall as a function of the spatial context within each experiment, we divided the images into 5 by 3 grids and calculated the summed frequency in which objects in each grid location were recalled (Figure 6). The total frequency for each location was captured in heat maps, where lighter shades indicate locations with greatest frequency and black indicates less frequency (note that some objects occupy more than one grid location, but were included in the frequency count in which most of the object occupied). We repeated this across each of the 3 context conditions.

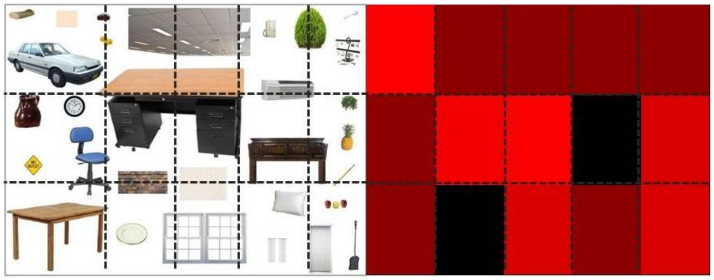
Figure 6 (top panel) shows a sample image and heat map from Experiment 1 (partial-context), the middle panel shows Experiment 2 (no-spatial-context), and the bottom panel shows Experiment 3 (random-context). First, the calculations for the heat maps revealed higher frequency for the number of objects recalled in the partial-context condition, relative to no-spatial and random conditions (166, 136, and 87, respectively). In addition, the most frequently recalled objects, in the partial-context image, cluster near the left and middle of the image– where that objects most iconic of an office appeared (i.e., desk, chair, computer, etc.), suggesting that memory is influenced by local contextual information. People most frequently recall objects that are associatively and spatially related (e.g. desk, chair, and computer). When the spatial relationships are



Partial Context Office



No Spatial Context Office



Random Context

Figure 6. Frequency heat maps with corresponding context scenes for all three experiments. Top panel: Partial-context scene (Experiment 1). Middle panel: No-spatial-context scene (Experiment 2). Bottom panel: Random-context scene (Experiment 3).

disrupted (middle panel), memory is better for the more salient (i.e. larger and strongly associated) objects (e.g. computer, desk, and chair) in the image and less centralized to a region of the scene. Notable, in the partial-context heat map, the lowest response region is across the top of the image, but in the no-spatial heat map, this is a high response region because it is occupied by the chair and computer.

The location of the most frequently recalled objects provide insight into the memory strategy participants adopted when recalling spatially disrupted scenes. It appears that people might use a ‘reading’ strategy when viewing a context disrupted image for memory (left to right, and across the top and down), with the most frequently recalled items at the top and left/middle of the image. When the associative relationships are disrupted, (bottom panel) memory performance further declines. In the random condition, memory appears to be better for objects salient in the image (e.g. car and desk).

2.5.4 Intrusions across contextual manipulations

The types of intrusions made further elucidated the contribution of global and local context to memory. For example, in the partial-context condition, participants incorrectly recalled objects that were consistent with the overall scene and prototypically accompany that scene type (e.g., when participants studied office scenes, they falsely recalled a calculator because this object is highly representative of objects generally found in office settings), suggesting that they were using global context information. Similarly, participants inaccurately recalled studying objects that are typically found in close proximity of some objects that were present in the scene (e.g., when participants studied urban scenes and saw a sky, they falsely recalled clouds because clouds are often found in the sky in urban settings), suggesting that they were using local context information.

3. Discussion

We presented an investigation assessing the relative contribution of global and local context to recall for scenes. First, we partially removed global scene context (i.e., background, walls, and ceiling) and found a decrement in performance for shorter study times. The removal of global context negatively affected the 2 second condition compared to the 10 and 15 second conditions. This might be due to the disrupted global level gist extraction with the removal of the background. Previous studies have found that global context supports quick interpretation of natural scenes (Potter, et al, 2002; Torralba, 2003), which appears to be important at shorter study times. During longer study times, available local context and additional study time may have compensated for this disruption. In the subsequent experiments, we systematically manipulated local context by removing spatial (i.e. randomize the object locations) and associative relationships of objects (i.e. randomize objects from various scene types). While the removal of the background of a scene initially impedes memory at short study times, the removal of spatial and associative context impinges on both short and long study times. When global and spatial information are removed performance at shorter study times is no better than guessing with prior knowledge (Original Study) after roughly 6 output positions. The results of the analysis for the number of responses further illustrate the benefit of global and local context, where the mean number of responses for full scene context > partial-context > no-spatial-context > random-context. The same trajectory applies to study time.

A critique of the current study is that it may not distinguish between whether context is truly assisting memory or if the benefit of context results from a participant bias to provide more responses. It is possible that having more contextual information available in the scenes encourages participants to make more guesses and although it may increase the number of correct responses, it may also increase the number of intrusions. However, this is unlikely, given that an increase in the number of intrusions would decrease the accuracy of responses and limit the difference in memory performance across the conditions where context is available. To rule out this possibility, in a follow-up study, participants may be forced to give a minimum number of responses across each of the contextual manipulations.

The findings from these experiments suggest that global and local context is important not only for scene perception, object recognition, and object categorization, but also for long-term retention in episodic memory. Furthermore, global and local context appear to make independent contributions to memory. Biederman (1982) asserted that spatial and global contexts are accessed simultaneously for the purposes of scene perception and object recognition. However, we found that accuracy decreased when objects were placed in scenes where the natural spatial position of objects were altered relative to when the objects were placed within consistent scenes and in contextual appropriate spatial positions. This suggests that although global and local spatial components of scene context make similar contributions, or at least are processed at the same rate and simultaneously in object and scene perception, these two features make substantively different contributions to long-term memory. While global context aids quick scene interpretation and memory at shorter study times, local context helps sustained memory accuracy at longer study times. This has important implications for understanding how the memory system works in natural environments, and for understanding failures of memory in situations that violate natural context. In particular, this raises tantalizing questions about the use to laboratory stimuli that is either disrupted or devoid of natural scene context.

Acknowledgments

This material is in review with the *Journal of Cognition*. It has been presented and published in the proceedings of the Annual conference of the Cognitive Science Society. The concept of this study was developed by K. Persaud and advisor, P. Hemmer. Stimulus creation, testing and data collection were performed by undergraduates, J. DeAngelis and R Venaglia, under the direct supervision of K.Persaud. Data analysis and interpretation were performed by K. Persaud and P. Hemmer. The thesis was drafted by K. Persaud and edited by advisor, P. Hemmer.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant Number NSF DGE 0937373, National Science Foundation I.G.E.R.T. under Grant Number NSF DGE 0549115, National Science Foundation REU under Grant Number IIS-1062735, and Rutgers University Aresty Summer Science Program.

References

Bar, M. (2004). Visual objects in context. *Nature Neuroscience*, 5, 617-629.

Bar, M. & Ullman, S. (1996). Spatial context in recognition. *Perception*, 25, 343-352.

Biederman, I. (1972). Perceiving real-world scenes. *Science*, *177,* 77-80.

Biederman, I., Mezzanotte, R.J., & Rabinowitz, J.C. (1982). Scene Perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143-177.

Brady T.F. & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science*, 19(7), 678-685

Galleguillos, C., & Belongie, S. (2010). Context based object categorization: A critical survey. *Computer Vision and Image Understanding*, 114(6), 712-722.

Griffiths, T.L. & Tenenbaum, J.B. (2006). Optimal predictions in everyday cognition. *Psychological Science*. 17(9), 767-773.

Hemmer, P. & Persaud, K. (2014). Interaction between categorical knowledge and episodic memory across domains. F*rontiers in Psychology*, 5(584)1-6.

Hemmer, P., & Steyvers, M. (2009). Integrating episodic and semantic information in memory for natural scenes. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Hemmer, P., and Steyvers, M. (2009b). A bayesian account of reconstructive memory. *Topics in Cognitive Science*, 1, 189–202.

Hemmer, P., Steyvers, M. & Miller, B. (2010). The Wisdom of Crowds with Informative Priors. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Portland, OR: Cognitive Science Society.

Hemmer, P., Tauber, S. & Steyvers, M. (in revision). Bayesian estimation in rational models.

Hollingworth A., & Henderson, J. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General 127*(4), 398-415

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in establishing spatial location. *Psychological Review*, 98, 352-376.

Huttenlocher, J., Hedges, L.V., & Vevea, J.L. (2000). Why Do Categories Affect Stimulus Judgment? *Journal of Experimental Psychology*, (129)220-241.

Inoue, K. & Takeda, Y. (2012). Scene-context effect in visual memory is independent of retention interval. *Japanese Psychological Research*, 54(4), 360-367.

Jern, A. & Kemp, C. (2013). A probabilistic account of exemplar and category generation. *Cognitive Psychology*. 66(1), 85-125.

Konkle, T., & Oliva, A. (2007). Normative representation of objects: Evidence for an ecological bias in object perception and memory. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society,* Nashville, TN*:* Cognitive Science Society.

Mandler, J.M & Johnson, N.S. (1976). Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 529-540.

Palmer, S. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*, 519 526.

Persaud, K., and Hemmer, P. (2014). The influence of knowledge and expectations for color on episodic memory. In P. Bello, M. Guarini, M. McShane, and B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, Quebec City, CA: Cognitive Science Society.

Potter, M. C., Staub, A., Rado, J., & O'Connor, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. Journal of Experimental Psychology: *Human Perception and Performance, 28*, 1163-1175.

Steyvers, M. & Hemmer, P. (2012). Reconstruction from memory in naturalistic environments. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation*, (pp.126-144). Elsevier Publishing.

Torralba, A. (2003). Contextual priming for object detection. *International Journal of Computer Vision, 53*, 169-191.