

How Visual and Semantic Information Influence Learning in Familiar Contexts

Annabelle Goujon

James R. Brockmole

Krista A. Ehinger

LPC-CNRS & Université
de Provence

University of
Notre Dame

Massachusetts Institute of
Technology

Previous research using the contextual cuing paradigm has revealed both quantitative and qualitative differences in learning depending on whether repeated contexts are defined by letter arrays or real-world scenes. To clarify the relative contributions of visual features and semantic information likely to account for such differences, the typical contextual cuing procedure was adapted to use meaningless but nevertheless visually complex images. The data in RT and in eye movements show that, like scenes, such repeated contexts can trigger large, stable and explicit cuing effects, and that those effects result from facilitated attentional guidance. Like simpler stimulus arrays, however, those effects were impaired by a sudden change of a repeating image's color scheme at the end of the learning phase (Experiment 1), or when the repeated images were presented in a different and unique color scheme across each presentation (Experiment 2). In both cases, search was driven by explicit memory. Collectively, these results suggest that semantic information is not required for conscious awareness of context-target covariation, but it plays a primary role in overcoming variability in specific features within familiar displays.

Humans navigate with ease in rich, complex and dynamic visual environments despite sparse and volatile representations of space (Irwin & Andrews, 1996; Rensink, 2000). To account for this apparent paradox, it is widely assumed that the visuo-cognitive system is adept at noticing and remembering regularities that remain constant in the environment (for a review, Oliva & Torralba, 2007), whether these be spatial or temporal, absolute or probabilistic (e.g. Baldwin, Anderson, Saffran, & Meyer, 2008; Chun & Jiang, 1998; Fiser & Aslin, 2001, 2002; Geng & Behrmann, 2005; Olson & Chun, 2001; Turk-Browne, Isola, Scholl, & Treat, 2008; Turk-Browne, Jungé, & Scholl, 2005; Turk-Browne, Scholl, Johnson, & Chun, 2010). By representing statistical contingencies present in the visual world, observers can reduce its cognitive complexity, which in turn can result in more efficient behaviour. In real-world scenes, such statistical contingencies can be based on either visual or semantic aspects of the environment. The present study aimed to clarify the respective contributions of these sources of information in supporting the learning and use of contingencies available in familiar contexts.

The ability to extract and use regularities based on visual and semantic contingencies may subservise different types of tasks. For example, visual contingencies might be more useful in processing highly familiar contexts, such as routine environments in which we navigate day-to-day while abstract semantic aspects of scenes would more likely be critical in allowing the transfer of knowledge to new situations (Brainerd & Reyna, 1990; Goldstone & Sakamoto, 2003; Rogers & McClelland, 2008).

The circumstances and situations in which visual and semantic learning occur can be framed in the context of differential learning effects that have been reported in the contextual cuing literature (Chun & Jiang, 1998; Brockmole & Henderson, 2006a). The contextual cuing paradigm is now a popular method to investigate the manner and degree to which human observers use statistical contingencies to facilitate visually guided behavior. The general principle of the contextual cuing paradigm consists of presenting regularities within search displays that predict a target's location and to expose participants to these regularities throughout the course of the task to determine if they can be learned. In the original version of the task (Chun & Jiang, 1998), participants are instructed to search for a T target among L distractors. Some displays are repeatedly presented while others are presented only once. Search times for the repeated configurations usually show a greater decrease over repetitions than search times for the unrepeated configurations. This benefit has been named by Chun and Jiang (1998) "contextual cueing".

Since its inception, the contextual cuing paradigm has been adapted for use in a wide range of visual contexts in order to study statistical learning and to characterize the generality of these learning principles (e.g. Brockmole, Hambrick, Windisch, & Henderson, 2008; Brockmole & Henderson, 2006a; Chua & Chun, 2003; Chun & Jiang, 1999; Endo & Takeda, 2004). Although the covariance between targets and specific contexts can be learned in very different types of display, both quantitative and qualitative differences have been observed depending on whether the repeated contexts are defined by simple stimulus arrays or by real-world scenes. Indeed, repeated contexts defined by color photographs are learned faster and give rise to greater benefits than contexts made up of letters (Brockmole & Henderson, 2006a). Scene-based contextual cuing also seems to rely heavily on global pattern analysis (Brockmole,

The authors thank Luis Jimenez and two anonymous reviewers for their helpful comments on this research. They also thank Deborah Cronin for her help with data collection. Correspondence should be addressed to Annabelle Goujon at LPC-CNRS & Université de Provence, Centre St Charles, Bât 9, Case D, 3 Place Victor Hugo, 13331 Marseille Cedex 3, France or to Annabelle.Goujon@univ-provence.fr.

Castelhano, & Henderson, 2006; Ehinger & Brockmole, 2008), categorical identity information (Brockmole & Henderson, 2006a; Brockmole & Vö, 2010), and conscious awareness of the predictive regularities (Brockmole & Henderson, 2006a), while local elements (Brady & Chun, 2007; Jiang & Wanger, 2004; Kunar, Flusberg, Horowitz, & Wolfe, 2007), surface features (Jiang & Song, 2005), and implicit learning mechanisms (Chun & Jiang, 1998; but see Smyth & Shanks, 2007) play a more prominent role in the development of memory for consistently arranged arrays of letters or abstract shapes. Because the contextual cuing paradigm reveals different types of learning behavior depending on the testing environment, one must question why such differences arise.

One hypothesis posits that the quality and extent of contextual learning is correlated with the absence or the presence of semantic information (Brockmole, Hambrick, Windisch, & Henderson, 2008; Brockmole & Henderson, 2006a, Brockmole & Vö, 2010). The use of semantic information to guide search to learned targets, when available, would have several consequences that fit well with the contrasts drawn above between scenes (semantically rich) and letter arrays (semantically impoverished). First, it would reduce the reliance on visual features and local statistical relationships. Instead, emphasis would be placed on the analysis of global patterns (Brockmole et al, 2006) and of categorical identity information (Brockmole & Vö, 2010; Goujon, 2011). Second, by enabling a dual-coding of visual properties and semantic labels, semantic information could facilitate the discrimination of different scenes and improve the recognition of familiar displays (Brockmole & Henderson, 2006a). Third, semantically rich displays could lead to conscious awareness of contextual regularities, which in turn, could enhance their consolidation and retrieval in memory (Brockmole & Vö, 2010; however, Goujon, 2011; Goujon, Didierjean, & Marmèche, 2007, 2009).

Direct evidence for a semantic influence in contextual cuing within scenes comes from a variety of observations. For example, when repeated scenes are inverted, which makes them more difficult to identify, the development of contextual cuing is slowed and its overall benefit reduced relative to upright scenes (Brockmole & Henderson, 2006a). In addition, contrary to simple displays (Brady & Chun, 2007), contextual cuing in scenes survives changes to local arrangements of objects, provided that the identity of the scene is not altered (Brockmole et al., 2006). Finally, contextual cuing is possible when scene categories predict target position, even if scene exemplars are never repeated. For example, contextual cuing effects have been demonstrated when targets were always located on pillows in (nonrepeated) bedroom scenes (Brockmole & Vö, 2010), or when multiple scene categories each predicted the (x,y) location of a target in space independent of object arrangement (Goujon, 2011). Nonetheless, the semantic hypothesis does not, at present, clearly predict conscious awareness of contextual regularities. While Brockmole and Vö

(2010) obtained evidence for explicit cuing in their study, Goujon observed implicit learning in hers, whether semantically rich displays were defined by real-world scenes (Goujon, 2011) or visually sparse displays composed of numbers (Goujon, Didierjean, & Marmèche, 2007) or words (Goujon, Didierjean, & Marmèche, 2009). Although the semantic regularities do not necessarily give rise to conscious awareness, whether and to what extent the presence of semantic information is really required for conscious awareness remains an open question.

An alternate hypothesis is that visual features, alone, may account for the differences observed between arrays of letters and real-world scenes. Contextual cuing generated by repeated contexts might merely depend on the quantity and quality of the visual cues present in a display. Real-world scenes are rich, dense, and heterogeneous in terms of visual features. In contrast, letter arrays are visually sparse and are highly homogeneous. While the differences in complexity observed across these stimulus types may be intuitive, they are nonetheless difficult to quantify. To illustrate our point and to motivate our experiments, in this report, we adopt the notion of “visual clutter” as a proxy measure of visual complexity although other constructs could also be used. In our opinions, visual clutter is a useful metric for measuring complexity because it reflects the variety of different visual features in a display and it has been shown to predict performance on various tasks including visual search (see Wolfe, 1998). There are several ways to define visual clutter quantitatively, and we illustrate three approaches in Figure 1 (see Rosenholtz, Li, & Nakano, 2007 for additional discussion and computational details). *Feature congestion* captures the local variability in visual features such as color, luminance, contrast, and orientation—features that are commonly used in modelling bottom-up aspects of overt attention allocation. *Subband entropy* provides a measure of information density by considering the amount of visual redundancy across a visual display. *Edge density* is a measure of subjective visual complexity based on the proportion of pixels in an image that define edges between surfaces. Using these measures, a more cluttered, and in our nomenclature a more complex image is expected to have greater feature congestion, entropy, and edge density. Indeed, by all three measures, it is apparent that real-world scenes are more complex than letter arrays which can lead one to ask what consequences these differences have on behavior.

The visual complexity of real-world scenes provides greater distinctiveness between individual displays, thereby improving learning and giving the observer sufficient information to lead to explicit recognition of the familiar patterns. Simply put, the differences observed between real-world scenes and simple displays might be due to the fact that more visual cues are generally available in real-world scenes which allow for better discrimination, recognition, and retention of individual search contexts. In favour of this hypothesis, Goujon et al (2007, 2009) observed implicit contextual cuing effects within displays that were semantically rich but visually

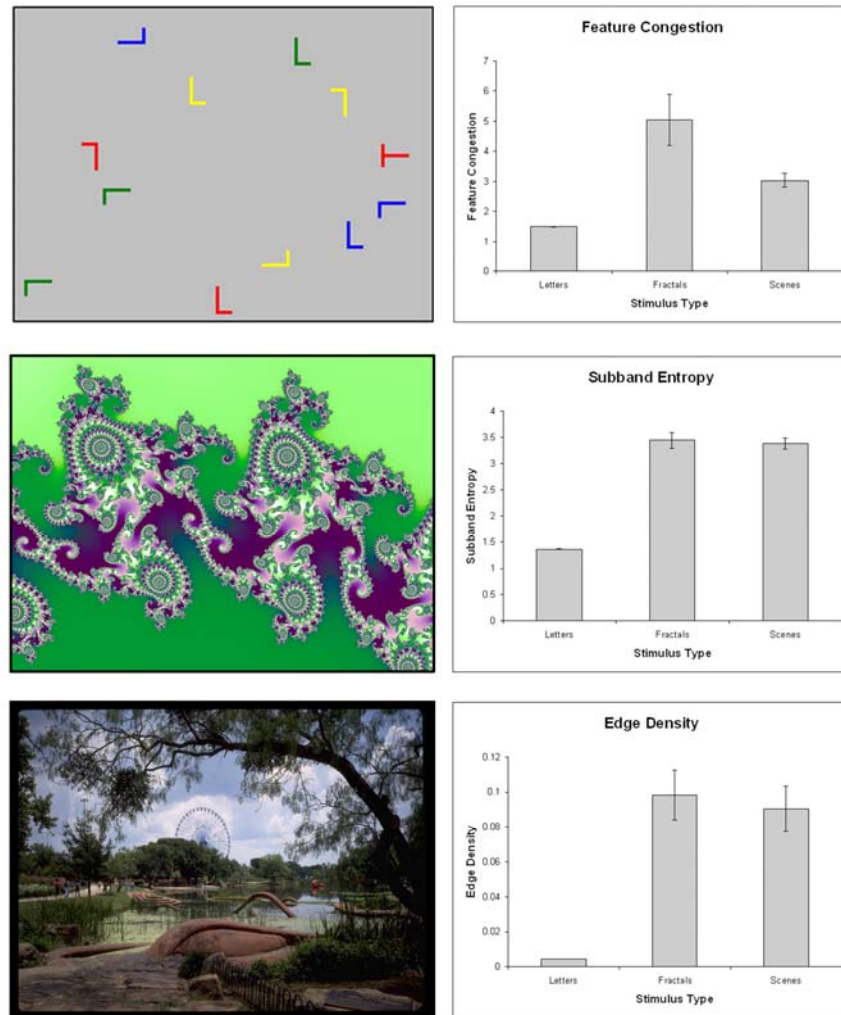


Figure 1: Three measures of visual clutter as a function of stimulus type (see Rosenholtz, Li, & Nakano, 2007). Data represents analysis of 8 letter arrays that were constructed according to the methods used by Chun & Jiang (1998), the 8 repeated fractals used in Experiments 1 and 2 of the current report, and the 8 repeated scenes used by Brockmole & Henderson (2006).

sparse such as spatial arrangements of numbers or words. These implicit effects emerged regardless of whether the regularities across displays were based on specific numbers or their category membership (i.e. when the evenness/oddness of the numbers in the context predicted the area where the target occurred in the display). In addition, the strength of those learning effects was closer to what is usually observed with repeated configurations of letters than with real-world scenes (for a discussion, see Goujon et al, 2009). Nevertheless, it remains possible that in real-world scenes, semantic information inherent to

certain visual features triggers different mechanisms that those that mediate learning in simple stimuli arrays.

Although contextual cuing has been explored in contexts both meaningless and meaningful and both simple and complex, these conditions have not yet been fully crossed. Specifically, a missing piece to the puzzle is search behaviour in displays that are visually complex but semantically sparse. This is a surprising oversight because the results obtained from such a circumstance can be used to directly contrast the semantic and visual hypotheses outlined above, and, by doing so, provide

much needed clarity pertaining to the relative contributions of visual features and semantic information in contextual cuing within repeated contexts and by extension in processing of familiar real-world scenes. Across two experiments, we used a contextual cuing procedure in which observers searched for letter targets embedded in images of fractals and other geometric patterns that were inherently meaningless but nevertheless visually complex. Indeed, the feature congestion, subband entropy, and edge density available in such stimuli can equal that in real-world scenes (see Figure 1). Hence, such images can mirror the complexity of scenes while lacking the semantic information available in scenes. If the behavioural differences between scenes and simple stimulus arrays are related to the semantic information present in a display, we would expect behaviour associated with fractals to resemble that observed with simple stimulus arrays. If, instead, the featural components of displays are sufficient to account for the differences between scenes and letter arrays, then we would expect behaviour associated with fractals to resemble that typically observed with scenes. In this report, we focused on two aspects of contextual cuing that have been shown to differ between scenes and simple stimulus arrays: the role of color information in learning, and the implicit or explicit nature of the acquired knowledge.

Divergent results concerning the importance of color information in contextual cuing have previously been observed in real-world scene contexts and in simple stimuli arrays. In scenes, the color content of search contexts has no influence on contextual cuing. Ehinger and Brockmole (2008) showed that modifying a scene's color scheme across each repetition affected neither the speed of learning nor the magnitude of the search benefit with respect to a control condition in which the scenes' color schemes were constant across repetitions. Moreover, a sudden and unexpected change to a scene's color scheme following multiple repetitions in a consistent color scheme did not impair contextual cuing performance. In simple stimulus arrays, however, changes to the color of search items can obliterate learning (Jiang & Song, 2005). On the one hand, this disparity may be explained by a semantic explanation in which scene identity, divorced from specific featural content, guides search. However, an alternative explanation based on visual features is just as feasible. In this case, the variability in color information is overcome by reliance on the multitude of other features present in the display that are unaffected by color changes (e.g., contours, edges, contrast, luminance, salience, , etc.). To contrast these hypotheses, the present experiments replicated the procedures used by Ehinger and Brockmole (2008), with visually complex but semantically sparse images. If semantic information is necessary to render color changes trivial, then contextual cuing in fractal images should be disrupted by color alterations. On the other hand, if multiple visual cues buffer against variability in a single feature, then color changes in fractals should be unaffected by color alterations.

In addition to examining the impact of color changes to contrast the semantic and featural hypotheses, we also examined whether the learning observed for fractal contexts was implicit or explicit in nature. To date, stimuli giving rise to explicit learning have been both semantically and featurally rich (i.e., repeated scenes), while implicit learning mechanisms underlie contextual cuing in featurally sparse displays (i.e., repeated arrays of letters or specific numbers displays). To access the implicit or explicit nature of learning that took place during our search task, participants were thereafter exposed to explicit memory tasks. Those tasks evaluated explicit knowledge related to both the identity of the repeated images and the target locations associated with those images. If these tasks reveal that contextual cuing in repeated complex and meaningless images results from explicit learning, this will show that semantic information is not required for conscious awareness of regularities.

Finally, we used this opportunity to enter into a recent debate concerning the use of response times (RTs) as a measure of attentional guidance in contextual cuing tasks. Some researchers have argued that reduced RTs indicate that contextual cuing facilitates the guidance of attention to targets (e.g. Brockmole & Henderson, 2006b; Chun & Jiang, 1998; Peterson & Kramer, 2001), others have argued that contextual cuing facilitates target processing after search is complete (Kunar, Flusberg, Horowitz, & Wolfe, 2007; Geyer, Zehetleitner, & Muleer, 2010). While analysis of RTs may not be able to resolve this debate, the recording and analysis of eye movements can because the eyes provide a direct and real-time measure of attention allocation. We therefore used both manual reaction times and eye movement behaviour as indicators of search efficiency.

EXPERIMENT 1

Experiment 1 had three aims: to characterise the temporal properties and search benefits associated with contextual cuing in visually complex but semantically poor images, to examine the role of color features in this cuing effect, and to determine the implicit or explicit nature of the involved learning mechanisms. To do this, the experiment was divided into two parts: a visual search task, immediately followed by explicit memory tasks.

In the first part of the experiment, participants searched for letter targets within images of fractals and other complex geometrical figures that were inherently meaningless but nevertheless visually rich and heterogeneous¹. Examples are illustrated in Figure 2. As with typical contextual cuing procedures, some of these images were repeatedly presented across blocks while others were shown only once during the task. This gave observers the opportunity to learn the position of the targets in some displays but not others. The occurrence of learning is revealed by a progressively increasing search benefit for repeated trials compared to novel trials.

¹ At debriefing none of the participants in this study indicated that they ascribed any inferred or personal meaning to the images.

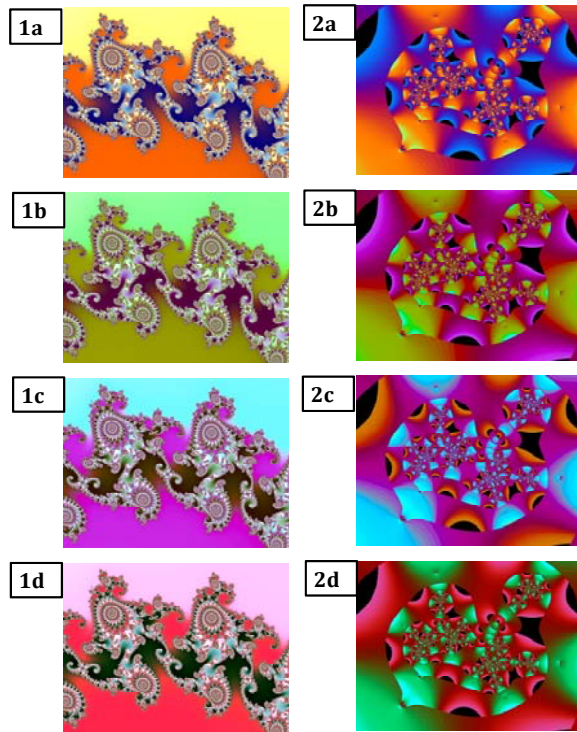


Figure 2: Representative stimuli used in the experiments. In Experiment 1, Panels 1a and 2a were used in the learning phase (blocks 1-15) and panels 1b and 2b were used in the transfer phase. In Experiment 2, all panels were used.

The role of color features in contextual cuing was investigated through a transfer block implemented at the end of the learning phase (Blocks 1-15). During this transfer block (i.e. Block 16), each repeated image was presented in an alternate color scheme. Past research has shown that a sudden alteration of color content has no effect on contextual cuing in real-world scenes (Ehinger & Brockmole, 2008) whereas a similar manipulation in simple stimulus arrays abolishes the benefit obtained in the previous blocks (Jiang & Song, 2005). If the density and the uniqueness of spatial features within our complex images are sufficient to overcome color modifications, the benefit obtained during learning should perfectly transfer without cost to the color-altered images as happens with scenes. In contrast, if the lack of color change effect observed in real-world scenes (Ehinger & Brockmole, 2008) was due, at least in part, to the presence of semantic cues, a cost would be observed during the transfer phase.

To determine the implicit or explicit nature of learning that took place during the search task, participants then completed a second experimental session which assessed their conscious memory for the images and target locations. To do this, observers were given a recognition task in which they identified images that were included in the search task. Then, they performed a target generation task on remembered images in which they

localized, to the best of their ability, the location of the target (which was not present at this stage).

Method

Participants. Ten individuals participated in Experiment 1. All reported normal or corrected-to-normal acuity with no color vision deficiencies. All participants received course credit.

Stimuli and Apparatus. The stimulus materials were based on 136 full-colour images of fractals and other abstract/geometrical patterns. 128 of these images were presented once during the experiment. The remaining 8 images were repeated throughout the experiment. From the original version of each repeated image, a single color-altered version was generated in CIE $L^*a^*b^*$ colorspace, following the procedures outlined by Ehinger and Brockmole (2008). This version was obtained by inverting the scene's a^* and b^* axes (equivalent to rotating the scene 180° in $L^*a^*b^*$ colorspace). As a result of this transformation, a color-opposite version of each image was produced which was maximally different in terms of color information. This color manipulation did not affect physical luminance levels (L^*) nor the distance within the color space between any given pair of pixels in the image. Within each image, a single gray “T” or “L” was presented in nine point Arial font. These letters constituted the target in a visual search task and were located in one of eight possible and equally occurring (x,y) coordinates. The luminance of the target was adjusted across images to approximately equate its contrast against the local background.

An additional eight images were created in the manner described above for a practice block of search trials. A third set of 16 images was created without targets for use in an explicit memory task (described below). All stimuli were displayed at a resolution of 800×600 pixels and in 24-bit color on a 19-inch screen with a refresh rate of 120Hz. Chin and forehead rests maintained a viewing distance of 80 cm. The position of each participant's right eye was monitored with an Eyelink 2K eye tracking system with a sampling rate of 1000 Hz and running in pupil and corneal-reflection mode, which resulted in an average spatial accuracy of 0.15° . A Microsoft video game controller was used to collect target classification responses.

Design and procedure. The experiment consisted of a visual search task immediately followed by two explicit memory tasks.

Search Task: Participants were instructed to search for the T or L in each image as quickly as possible, and to classify the target by pressing one of two buttons on the video game controller. The sequence of experimental trials was divided into 15 learning blocks (Blocks 1-15) and one transfer block (Block 16). Within each block, the eight repeated images were each presented once. The remaining eight trials presented unrepeated images that

were shown once during the experiment. The same subset of unrepeatd images occurred within the same block for all subjects. However, the order of trials within a block as well as the order of blocks was randomized for each subject. Within each repeated image, the target's location was held constant across presentations but its identity was randomly determined, with the constraint that each target appeared equally often. T and L targets also occurred an equal number of times across the unrepeatd trials as did the eight possible target locations. Critically, in Block 16, each repeated scene was displayed in its color-opposite version.

Participants began the search task by completing a calibration routine that served to map the output of the eyetracker onto display position. Calibration was monitored throughout the experiment and was adjusted when necessary. Participants then completed a practice block of eight trials to familiarize them with the procedure. The participants then moved seamlessly on to the experimental trials. At the beginning of each trial, a blue dot was centred on a white background. Observers were instructed to look at this dot and to press a button when ready to view the next image. Participants had to search as quickly as possible for the target ("L" or "T") and press the corresponding key for that target when it had been identified. The trial was terminated if a response was not made within 20 s of scene onset. Subjects were offered the opportunity for a short break after completing 128 trials, however, the trials were self paced and participants were informed that they could take additional breaks as they pleased.

Explicit memory tasks: After completing the search task, observers completed two explicit memory tasks to test their awareness of the repeated configurations. They were exposed to 32 trials which separately presented the eight repeated images from the search task (displayed in the color- version used during the transfer phase²), eight unrepeatd images used in the search task, and 16 new images that were not included in the search task. Importantly, none of these images contained a target. For each image, observers made two judgments. First, they indicated whether or not they remembered seeing the image during the search task (image recognition task). Then, they used a mouse trackball, to move a dot to the location they remembered the target being presented (target generation task).

Results and Discussion

Search Task. Analysis of the search task focused on two behavioural measures of search efficiency. The first, and more traditional measure, was overt response time, or the elapsed time between the start of the trial and the observer's response. We also considered the elapsed time between the onset of the trial and the observer's first fixation on the target (or more precisely, the first fixation within 50 pixels of the target). This second approach

² The consistent color versions used during the learning phase were used in the explicit memory tasks of Experiment 2

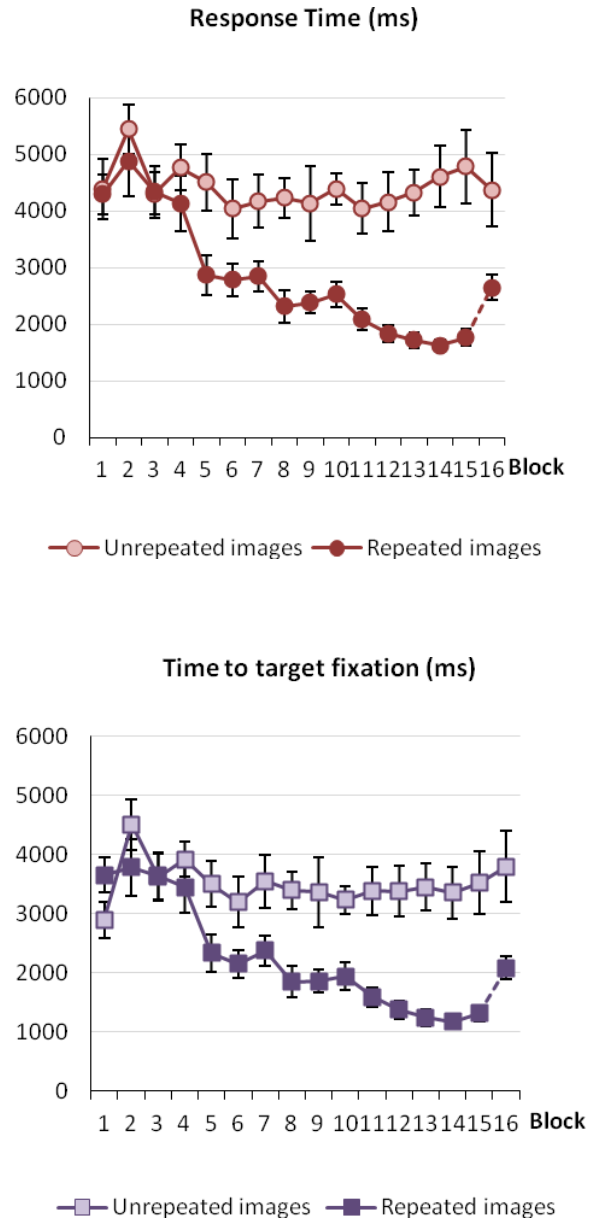


Figure 3: Mean response times (Panel a) and mean elapsed time prior to target fixation (Panel b) in each block for the repeated and unrepeatd conditions in Experiment 1. Block 16 corresponds to the transfer in the color schemes. The error bars show the standard error of the mean (N = 10).

enabled us to remove non-search components of RT such as the time needed to determine the target's identity and to execute a motor response from the dependent measures. Using gaze also has a theoretical advantage over RT because it provides a direct measure of attention guidance that cannot be ascertained from manual responses.

Trials were excluded from analyses if a response was not made within 20 s (0.8% of repeated trials & 6.2% of

unrepeated trials), was incorrect (2.0% of repeated trials & 0.9% of unrepeated trials), if the RT was greater than 3 standard deviations from the mean of unrepeated trials computed on a subject-by-subject basis (2.7% of repeated trials and 3.8% of unrepeated trials) or if no fixation was recorded within the interest area defined around the target (2.5% of the remaining repeated trials after exclusion of the errors and outliers and 4.2% of the remaining unrepeated trials). Response time and the time to target fixation are presented in Figure 3 as a function of condition and block.

Learning phase: A first series of analyses was aimed at characterizing the trends in search efficiency across the learning blocks (1-15). A repeated-measures analysis of variance (ANOVA) was conducted on each dependant variable (RT and time to fixation) with condition (repeated images vs. unrepeated images) and block (1-15) as within-subject factors. The ANOVAs showed main effects of block on both dependant variables: [RT: $F(14, 126) = 6.29, p < .001$, time to target fixation: $F(14, 126) = 6.34, p < .001$], as search became overall more efficient as the experiment progressed. Main effects of condition were also observed: [RT: $F(1, 9) = 46.38, p < .001$, time to target fixation: $F(1, 9) = 38.08, p < .001$], as search was more efficient in the repeated condition compared to the unrepeated condition. Finally, block by condition interactions were apparent: [RT: $F(14, 126) = 3.86, p < .001$, time to target fixation: $F(14, 126) = 3.70, p < .001$], indicating a progressively increasing benefit to both RT and the elapsed time prior to fixating the target in the repeated condition relative to the unrepeated condition.

The contextual cuing effect (i.e. difference between repeated an unrepeated trials) was significant from Block 5, that is, after 4 repetitions and remained robust throughout the remainder of the task. Trends in both dependant variables were entirely consistent, suggesting that the contextual cuing effect resulted from a guidance benefit. Ultimately, learning produced a final benefit of around 2700ms in RT and 2200ms in elapsed time prior to fixating the target. Thus, as real-world scenes (e.g. Ehinger & Brockmole, 2008), complex meaningless images triggered large and stable contextual cuing effect throughout the course the task.

Transfer phase: A second series of analyses investigated the impact of the color change on search performance using paired samples *t*-tests. Critically, these analyses revealed that the RT and the elapsed time to target fixation observed in the repeated condition were both longer in the transfer block (Block 16) than in the final learning block (Block 15) [$t(10) = 3.77$ and $t(10) = 3.48$ respectively, all $ps < .005$]. Therefore, the color change impaired the contextual cuing effect generated at the end of the learning phase. The cost averaged 950ms in RT and 850ms in elapsed time to target fixation. However, this cost was not complete; a significant benefit (i.e. 1700ms in both RT and elapsed time to target fixation) was nevertheless obtained on both dependant variables in the repeated condition compared to the

unrepeated condition in Block 16, $t(10) = 2.59$ and $t(10) = 2.72$ respectively, all $ps < .05$.

The results of the transfer phase gave rise to two main conclusions. First, contrary to the pattern of behavior observed in real-world scenes (Ehinger & Brockmole, 2008), a sudden change in a meaningless complex image's color scheme impairs contextual cuing. This cost was observed in each participant (ranging from 300ms to 1800ms in RT) and for each of the eight color-altered repeated images (ranging from 300ms to 2300ms in RT). This result is consistent with the hypothesis that in real-world scenes, semantic information is used to overcome variability in specific visual features that occur between individual encounters with a visual display. Second, because these costs were observed in both dependent measures, they cannot be ascribed solely to the response selection phase of search. Indeed, comparing the average observed cost in RT (950 ms) to the average cost in the time needed to fixation the target (850 ms), it is apparent that the majority of the cost was generated during the search itself.

Explicit memory tasks

Image Recognition Task. Participants reported remembering 96.6% of the images that were repeated during the search task, 52.5% of those that were presented once during the search task (i.e. belonging to the unrepeated condition in the search task), and 17.5% of the new images never seen during the search task (corresponding to false alarms). The recognition ability was then measured by calculating the discrimination indexes (A') per participant for both the repeated and unrepeated conditions from the hit (H) and false alarms (FA) rates (see Stanislaw & Todorov, 1999). Comparisons to chance (.5) showed that performance was reliably above chance level for both the repeated ($M=.94, SE=.02$) [$t(9)=24.09, p<.001$] and unrepeated ($M=.74, SE=.05$) [$t(9)=4.35, p<.005$] conditions. Comparison between those A' values indicated that they differed significantly [$t(9)=4.49, p<.005$].

Target Generation Task: To evaluate the accuracy of observers' memory for target locations, the distance between the actual target location and the observer's placement of the target was measured (Brockmole & Henderson, 2006a). For the 16 new images, an arbitrary referencing location was determined from the eight coordinates used in the search task. The mean distance error obtained in the new condition gave an indicator of random responses. The average placement error was 135 pixels for the repeated images, 326 pixels for the novel images and 369 pixels for the new images, that is, 5.85°, 14.10°, and 15.98° of the visual angle respectively. A one way repeated-measures ANOVA showed a main effect of condition [$F(2, 18) = 37.3, p < .001$]. Pairwise comparisons indicated that the repeated condition differed significantly from the unrepeated and new conditions, all $ps < .001$, but that the unrepeated and new

conditions did not differ from each other, [$t(9)=1.16$, $p=.141$].

The results of the explicit memory tasks indicate that knowledge underlying the contextual cuing effect observed at the end of the search task was explicit. The image recognition rates were highly accurate for the repeated images as was memory for target locations within the repeated images. In addition to these measures, at debriefing, all participants reported having noticed that certain images were repeated, that the target location was held constant within these images and that this knowledge helped them in their searching of the target. They all spontaneously mentioned the color change, but contrary to the behavioural evidence, most had the feeling that the color alteration did not impair their search performance. Therefore, semantic information is not necessary for conscious awareness of learned regularities.

EXPERIMENT 2

Experiment 2 sought converging evidence for the conclusions drawn in Experiment 1. In Experiment 1, color may have been used as a reliable cue for target position because it remained constant throughout the learning phase of the experiment. Hence, when it was altered, an associated cost in search was observed. In Experiment 2, we eliminated any possibility that color information could be used to reliably reference target position. To do this we compared learning in two situations. First, repeated images were presented in the same color scheme with each repetition. Second, repeated images were presented in a different and unique color scheme with each repetition. If color, without semantic support, is an important feature when developing context-target associations, then learning rates should be slower in the variable color condition.

Method

Participants. Thirty eight individuals participated in Experiment 2. All reported normal or corrected-to-normal acuity with no color vision deficiencies. None had participated in Experiment 1.

Stimuli: The stimuli used in Experiment 1 were also used in Experiment 2. In addition to their original version (used during the learning phase of Experiment 1), 15 unique alternate color versions of each repeated image were also generated. Seven transformed images were initially produced using the axis swap (e.g. changing yellow to red) and axis invert (e.g. changing yellow to blue) operations described by Oliva and Schyns (2000). These seven altered versions and the original version of the scene were then further transformed by rotating the color axes 45° in L^*a^*b color space (see Ehinger & Brockmole, 2008, Experiment 1). Where necessary, the levels of cyan, magenta or yellow present in an image were manually adjusted with Adobe Photoshop to subjectively maximize the differences between the 16 versions of each repeated image.

Apparatus, design and procedure: The apparatus, the design and the procedure were the same as those in Experiment 1, unless otherwise indicated. Participants were randomly assigned to one of two groups that differed with regard to the color schemes of the stimuli. All the participants were exposed to a sequence of 16 blocks of 16 trials. For those in the consistent color group, repeated images were presented in 1, and only 1, of its 16 versions, i.e. in the original colored version used in learning phase of Experiment 1. For those in the variable color group, repeated images were presented in all 16 possible color versions across the experiment. In both groups, the unrepeated images were identical to those used in Experiment 1. The stimuli and the procedure used in the explicit memory task were the same as in Experiment 1, except that the repeated images were presented in their original colored version.

Results and discussion

Trials were excluded from analyses if a response was not made within 20 s (1.1% of repeated trials & 4.0% of unrepeated trials), was incorrect (1.3% of repeated trials & 2.6% of unrepeated trials), if the RT was greater than 3 standard deviations from the mean of unrepeated trials computed on a subject-by-subject basis (1.2% of repeated trials and 3.6% of unrepeated trials) or if no fixation was recorded within the interested area defined around the target (3.4% of the remaining repeated trials after exclusion of the errors and outliers and 4.8% of the remaining unrepeated trials). Response times and the elapsed time prior to target fixation are presented in Figure 4 as a function of the experimental factors.

Because we were to compare performance across two groups of participants, we initially conducted an analysis that strictly focused on the unrepeated trials. These trials establish a baseline search speed which needs to be ascertained before group comparisons can be carried out. We therefore conducted a mixed-model ANOVA crossing the within-subject factor of block (1-16) and the between subjects factor of group (consistent color vs. variable color). Although both groups viewed exactly the same unrepeated images, unknown differences between the groups resulted in a reliable main effect of group in RT [$F(1,36) = 5.26$, $p < .05$] and a marginally reliable effect in time to target fixation [$F(1,36) = 3.45$, $p = .07$]. In both measures, participants belonging to the consistent color group took less time to complete their searches on novel trials than the participants belonging to the variable color group, i.e. 3513 ms vs. 4081 ms in averaged RT and 2853 ms vs. 3227ms in averaged time to target. The block effect was not significant, in RT [$F(3,108) = 1.28$, $p = .206$] and in time to target fixation, [$F(3,108) = 1.20$, $p = .263$]. The interaction of these factors was not reliable, both $F_s < 1$.

To account for the differences in baseline search rates across groups in our analyses of learning, we focused on a proportional measure of performance (a *cuing effect*) where the difference in search time between the

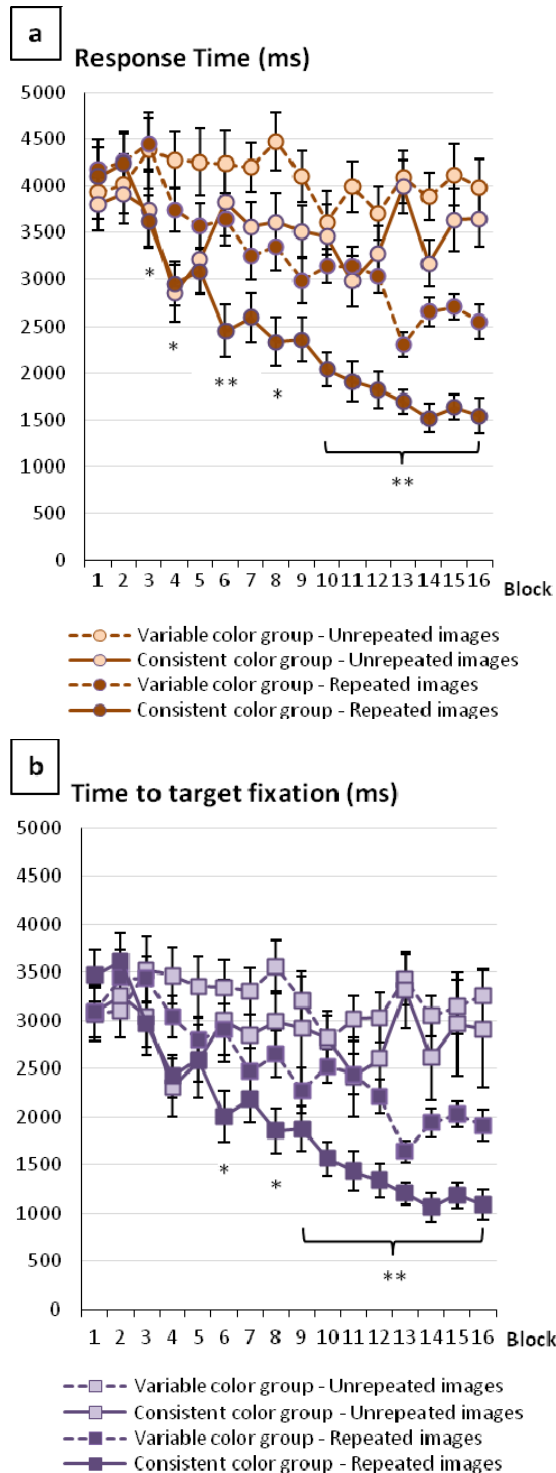


Figure 4: Mean response times (Panel a) and mean elapsed time prior to target fixation (Panel b) in each block for the repeated and unrepeated conditions and in each consistent and variable color conditions in Experiment 2. The error bars show the standard error of the mean ($N = 19$). The asterisks indicate the relevant differences between the repeated consistent vs variable color conditions (* = $p < .05$ and ** = $p < .005$, t -tests).

unrepeated and repeated trials was divided by the search performance observed for unrepeated trials (see Brockmole et al, 2008; Davoli, Brockmole, & Goujon, 2011). Using this procedure, for example, a cuing effect of .33 would indicate that search performance for repeated trials was one-third faster than unrepeated trials. Contextual cuing effects based on the response times and the elapsed time prior to target fixation are illustrated in Figure 5 for each group as a function of block.

We conducted a mixed-model ANOVA in which proportional measures of RT and time to target fixation were considered as a function of group (consistent vs variable) and block (1-16). A main effect of block was observed in both dependent measures [RT: $F(15,540) = 15.06$, $p < .001$; Time before target fixation: $F(15,540) = 11.64$, $p < .001$]. Additionally, a main effect of group was observed in RT [$F(1,36) = 6.78$, $p < .05$] and a marginal effect of group was observed in time to target fixation [$F(1,36) = 3.20$, $p = .082$]. Critically, both measures revealed a reliable interaction between these factors [RT: $F(15,540) = 1.88$, $p < .05$; Time before target fixation: $F(15,540) = 2.35$, $p < .005$]. This interaction arose because contextual cuing was weaker in the variable color condition relative to the consistent color condition. From this observation, we can conclude that contextual cuing in repeated complex meaningless pictures was affected by the nature of the color content. The complexity of spatial visual features characterizing the present materials was not sufficient to fully overcome the color alterations, even when color constituted an irrelevant cue for learning.

Explicit memory tasks. One participant did not complete the explicit memory task. The analyses were therefore conducted on 37 participants.

Recognition task: Image recognition among members of each group (consistent color and variable color) as a function of trial condition (repeated, unrepeated and new), as well as discrimination indexes (A') calculated per participant in each group for both the repeated and unrepeated conditions are summarized in Table 2. As in Experiment 1, comparisons of those discrimination values to chance (.5) showed that performance was reliably above chance level in each group and for both conditions, all $p < .001$. Comparison of the A' values between both groups and chance showed a significant difference in the repeated condition [$t(35) = 2.97$, $p < .01$], but not in the unrepeated condition, [$t(35) = 1.12$, $p = .27$]. Those results suggest that color manipulations during the search task had an impact on the recognition of the familiar images, by creating confusion between the exemplars used in the experiment.

Target generation task: A mixed-model ANOVA crossing the within-subject factors of condition (repeated, unrepeated vs. new), and the between subjects factor of group (consistent color vs. variable color) was also conducted on the placement errors in the target generation task. These data are also summarized in Table 2. This

Table 1: Image recognition and Placement errors among members of each group (consistent color and variable color) as a function of trial condition (repeated, unrepeated and new). The placement errors are reported both in pixels and in degrees of visual angle.

Task	Group	Condition		
		<i>Repeated</i>	<i>Unrepeated</i>	<i>New</i>
<i>Percentage of images recognized as "seen"</i>	Consistent color	97.3% (2%)	48.7% (4%)	12.8% (7%)
	Variable color	92.4% (2%)	64.6% (4%)	20.1%(2%)
<i>Discrimination Index (A')</i>	Consistent color	.96(.01)	.78(.02)	
	Variable color	.92(.01)	.81(.02)	
<i>Placement error in Target Generation Task</i>	Consistent color	115 px (24) =5.00°(1.06)	388 px (14) =16.78°(0.61)	377 px (10) =16.29°(.05)
	Variable color	194 px (25) =8.38°(1.08)	368 px (14) =15.93°(0.62)	345 px (11) =15.94°(0.48)

analysis showed a main effect of trial condition [$F(2,70) = 102.79, p < .001$], no effect of group, $F(1,35) < 1$ and a condition x group interaction, [$F(2,70) = 6.00, p < .005$]. Post hoc analyses indicated that in both consistent color and variable color groups, the distance between the actual target location and the observer's placement of the target was significantly smaller in the repeated images than in the unrepeated and new images. No difference between the unrepeated and new conditions was observed. Crucially, the placement error in the repeated images was reliably greater in the variable color group than in the consistent color group, suggesting that the modifications of the color schemes in the search task had an impact in the generation task as well.

Combined with the recognition task, the results of the target generation task confirm the explicit nature of the learning mechanisms involved in contextual cuing within complex and meaningless images. Despite alterations of the color schemes, contextual regularities were explicitly detected and learned by the participants. However, and although the hit rates for the repeated images did not differ between the variable color and consistent color groups [$t(35)=1.73, p=.093$], the A' discrimination index was higher in the consistent color group than in the variable color group. Thus, even if the task did not reveal the time of response selection processes, the color manipulations had an impact on the recognition of images. In addition, the participants' ability to accurately place the target was reliably worse among those exposed to the variable color version of the materials in the search task (with an error of 194 px, i.e. 8.4° of the visual angle) compared to those exposed to the consistent color version (115 px, i.e. 5.0° of the visual angle). This result supports a role of color features in learning context-target location associations. Therefore, the color features were not merely involved in recognition of familiar visual patterns, but were also included in precision of context-target representations.

General Discussion

Past investigations into the information used to support behaviour in familiar contexts has considered a variety of visual displays. Those investigations highlighted important differences in search behaviour, depending on whether the familiar contexts were defined by arrays of letters and other simple shapes (e.g. Chun & Jiang, 1998) or composed of real-world scenes (e.g. Brockmole & Henderson, 2006a). Because these classes of stimuli differ in both their visual complexity and semantic richness, the source of these learning differences is unclear. While previous work has divorced these aspects of an image by using arrays of categorically related words or numbers to generate displays of low visual complexity and high semantic richness (Goujon et al., 2007, 2009), the present study was the first to explore learning and search behaviors in stimuli that are highly rich in terms of visual cues, but poor in terms of semantic cues. In two experiments, observers searched for a target letter in fractals and other complex geometric patterns. A subset of these images was repeated so that observers could potentially learn the covariance between images and target placements. Two aspects of search efficiency were examined: The role of feature information (specifically color) in such learning, and the explicit versus implicit nature of the acquired knowledge.

The results of this study revealed that visually complex and meaningless images can trigger large, stable and explicit contextual cuing effects similar to those found with real-world scenes, and that those effects result from facilitated attentional guidance (for a summary of the effects, see Figure6). That said, the overall pattern of learning obtained in Experiments 1 and 2 aligned more with the pattern of results usually observed with simple stimulus arrays rather than with real-world scenes in that color manipulations had a negative impact on search. Experiment 1 showed that a sudden change of a repeating

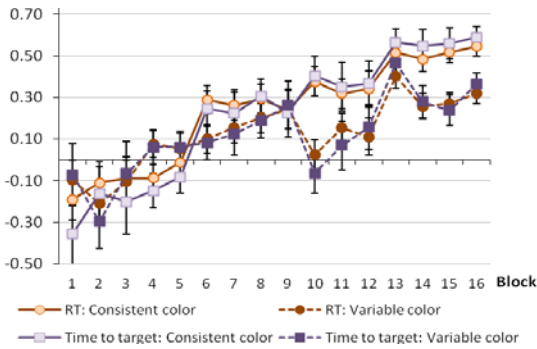


Figure 5: Cuing effect in response times and elapsed time prior to target fixation as a function of block for the consistent and variable color groups. The error bars show the standard error of the mean ($N = 19$).

image's color scheme impaired the contextual cuing effect produced over previous encounters with the display. In Experiment 2, target locations within repeated images that were presented in a different and unique color scheme across each presentation were learned more slowly than locations within images that were presented in the same color scheme with each presentation. This result is confirmed by lower performance in the explicit memory tasks. Therefore, visual complexity, at least defined by high feature congestion, high subband entropy and high edge density, was not sufficient to fully overcome color alterations. This suggests that semantic information may play a primary role in overcoming variability in specific features within familiar displays. This may occur because semantic regularities could give the visuo-cognitive system more flexibility by enabling a dual-coding of visual properties and semantic labels. This in turn may result in visual representations that are less specific and less dependent on visual features, even in familiar contexts. In addition, semantic information could also improve the frame-of-reference and the target placement within the scenes.

Although semantic information may be an important guide for search, it is not required for conscious awareness of context-target covariations. Indeed, in both experiments, participants were able to reliably identify the repeated images as well as the target locations. The high recognition rates for the repeated images (97% vs. 12% of false alarms for the consistent color group in Experiment 2) and the accuracy in target placement (5.0° in repeated images vs. 16° of the visual angle in new images), closely resemble that observed in real-world scene viewing (97% vs. 9% in a recognition task and 1.7 cm (2.5°) vs. 9.3 cm (13.3°) in a target generation task, see Brockmole & Henderson, 2006a). On the other hand, this pattern of performance contrasts with the implicit cuing effects observed in simple stimuli arrays, whether the predictive

regularities are based on repeated configurations of letters (Chun & Jiang, 2003), specific numbers displays (Goujon et al., 2007) or semantic category information (Goujon et al., 2007, 2009).

Therefore, semantic cues are not required to account for conscious awareness associated to learning of contextual regularities in real-world scenes. Instead, visual uniqueness of images, due to their complexity and to the visual heterogeneity between different exemplars, seems to be sufficient to allow the observer to consciously recognize and discriminate repeated patterns. Of course, our data do not exclude the possibility that semantic information can trigger explicit memory and it would probably be unreasonable to claim that explicit memory is only triggered by visual aspects of an image. Instead, we think it is likely that any factor that increases the distinctiveness between visual environments would increase the likelihood that explicit knowledge would emerge. This interpretation may explain why implicit as well as explicit semantic contextual cuing effects may have been reported in the literature (e.g. Brockmole & Vö, 2010; Goujon, 2011). Furthermore, we suspect that while a scrambled scene might retain the same visual characteristics of an unadulterated scene, the loss of semantic coherence or of visual internal structure would disrupt explicit memory.

The third contribution of this study concerns the trends of learning within repeated meaningless but complex displays and the nature of the involved mechanisms. Although our results highlight a contribution of semantic information in contextual cuing within repeated real-world scenes, the strength of the effects observed with the present materials shows that complex, heterogeneous but meaningless images can trigger cuing benefits as large and stable as those found in scenes. However, the speed of learning in complex meaningless pictures was more similar to the speed usually observed with simple displays. In the consistent color conditions of Experiments 1 and 2, the contextual cuing effect was significant after four or five repetitions, whereas only two or three repetitions are usually required with scenes (e.g. Brockmole & Henderson, 2006a, Ehinger & Brockmole, 2008). In line with a result reported by Brockmole and Henderson (2006a, Experiment 2), semantic cues present in real-world scenes may speed the acquisition of learning that underlies contextual cuing. Those semantic cues could provide the observer with extra information and enable dual-coding of this information, which could enhance the detection of contextual regularities.

The question was then to determine how learning of the contextual regularities had facilitated the search task. The eye movement data in both experiments suggests that repeated complex meaningless images, as real-world scenes, had facilitated search mainly by orientating attention toward the target location. Indeed, an important and growing benefit in elapsed time prior to fixating the target was observed in repeating contexts, and this benefit was entirely correlated with the benefit found in RT. The good performance in the generation task supports the conclusion that participants acquired strong knowledge

Study	Characteristics of the materials				Characteristics of the contextual cuing		
	Contexts	Regularities based on	Subjective visual complexity of the displays	Semantic information	Nature of the regularities	Nature of knowledge	Impact of color manipulations
Chun & Jiang (1998) Jiang & Song (2005)	Ls	Repetition of spatial configurations	impoverished	Absent	Visual	Implicit	Yes
Goujon et al (2007), E1	Numbers	Repetition of specific numbers	impoverished	Present	Visual & Semantic	Implicit	
Goujon et al (2007), E2, E3 Goujon et al (2009)	Numbers Words	Evenness/oddness property Semantic category	impoverished	Present	Semantic	Implicit	
Brockmole & Henderson (2006a) Ehinger & Brockmole (2008)	Real-world scenes	Repetition of scenes	Dense & heterogeneous	Present	Visual & Semantic	Explicit	No
Brockmole & Vo (2010)	Bedroom scenes	Object containing the target	Dense & heterogeneous	Present	Semantic	Explicit	
Goujon (2011)	Indoor scenes	Semantic category	Dense	Present	Semantic	Implicit	
Present study	Fractal images	Repetition of images	Dense & heterogeneous	Absent	Visual	Explicit	Yes

Figure 6. Characteristics of the contextual cuing effects depending of the characteristics of the testing environments

regarding the images-target location associations. Although a similar interpretation had been initially proposed by Chun and Jiang (1998) to explain contextual cuing in the classic task, this latter one is currently debated. Because contextual cuing in simple stimuli arrays relies more upon the repetition of the local context than upon the global context, an explanation in terms of guidance benefit seems to poorly account for contextual cuing effects observed with the classic task (Kunar et al, 2007, 2008; however, Junge & Chun, 2007; Fisk & Sanocki, in press). Such effects might, in some part, result from enhanced feature contrast signals at the overall-saliency computation stage (Geyer, Zehetleitner, & Muller, 2010) or from facilitating response selection (Kunar et al, 2007).

The further question is then to explain why certain testing environments exert a strong impact on attentional guidance whereas others have a weak influence. A fundamental difference between the effects reported in the present paper and those observed with the classic task concerns the conscious awareness of the contextual regularities. It is likely that conscious awareness of regularities enhances their consolidation and retrieval in memory but also favours deployment of attention toward the target location (see Kunar et al, 2006 for a similar argument). During a visual search, several factors and mechanisms (top-down or bottom-up driven) compete to control the guidance of selective attention. A memory-based search requires first the retrieval of context-target representation and then the interaction of that memory trace with attentional guidance (Brockmole & Henderson, 2006b; Peterson & Kramer, 2001). It is possible that when the knowledge is not accessible to awareness, a more competitive mechanism “wins the race” and consequently leads to frequent “memory failures” (Logan, 1988; 2002). This might explain the weak impact of implicit contextual cuing on attentional guidance. By providing subjective coherence between a situation and an adapted behavior, consciousness might force memory-based performance.

Thus, conscious awareness of regularities might explain the differences in the strength of the contextual cuing effects (see Goujon, 2011 for a similar argument). This argument highlights the adaptive role of consciousness in analysis of visual scenes.

Placed in a broader context, the results of this study also raise new questions related to the processing of real-world scenes. Previous research has shown that visual clutter impairs performance in visual search and that it may in fact serve as an image-based proxy for set size in real-world scenes (Rosenholtz et al, 2007; Wolfe, 1998; Henderson, Chanceaux, & Smith, 2009). The present study, however, suggests that increased levels of visual clutter may enhance learning and consequently improve memory-based search. Although clutter impairs bottom-up search by saliency or target features – because the variety of features in a cluttered scene compete for attention – the relationship between clutter and complexity means that clutter may actually facilitate top-down guidance by contextual regularities and memory. Clutter does not, itself, facilitate top-down search, but the greater variety of features in complex scenes means that there is more opportunity for unique configurations, which can be used to distinguish scenes and support associations between scene context and target locations. This shows that relationships between visual clutter and visual search performance are more complex than a mere linear correlation. In addition, the semantic influences highlighted in our study reinforces arguments that while modeling attentional control through bottom up-processes such as saliency and local image statistics is important, it also provides an incomplete characterization of behavioral mechanisms (Henderson, Brockmole, Castelhana, & Mack, 2007; Torralba et al., 2006). More work investigating the interactions between visual and semantic aspects of attentional control is still needed.

To conclude, the present study offers new findings concerning the respective contributions of semantic information and visual features in learning and using of

contextual regularities present in repeated complex visual environments. Surprisingly, meaningless but visually complex images tend to trigger learning mechanisms producing similar search benefit than real-world scenes. Featural components are also sufficient to trigger explicit learning effects, suggesting that semantic information does not play a necessary role in attaining conscious awareness of scenes-target associations. However, by enabling a dual coding of the information, semantic information might speed the detection of regularities and allow the observers to overcome variability and visual degradations in the learned context, such as color modifications. This may be because semantic cues reduce the reliance on visual features and on their statistical relationships.

REFERENCES

- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*, 1382-1407.
- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 798-815.
- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Research*, *10*, 3-47.
- Brockmole, J. R., Castelano, M. S., & Henderson, J. M. (2006). Contextual cuing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 699-706.
- Brockmole, J. R., Hambrick, D. Z., Windisch, D. J., & Henderson, J. M. (2008). The role of meaning in contextual cueing: Evidence from chess expertise. *Quarterly Journal of Experimental Psychology*, *61*, 1886-1896.
- Brockmole, J. R., & Henderson, J. M. (2006a). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*, 99-108.
- Brockmole, J. R., & Henderson, J. M. (2006b). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 1177-1187.
- Brockmole, J. R., & Vö, M. L-H. (2010). Contextual cuing without repeated episodic experience: semantic memory has an independent role in learning contextual regularities in real-world scenes. *Attention, Perception & Psychophysics*, *72*, 1803-1813.
- Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics*, *65*, 72-80.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28-71.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*, 360-365.
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 224-234.
- Davoli, C. C., Brockmole, J. R., & Goujon, A. (in press). A bias to detail: How hand position modulates visual learning and visual memory. *Memory & Cognition*.
- Ehinger, K. A., & Brockmole, J. R. (2008). The role of color in visual search in real-world scenes: Evidence from contextual cueing. *Perception & Psychophysics*, *70*, 1366-1378.
- Endo, N., & Takeda, Y. (2005). Use of spatial context is restricted by relative position in implicit learning. *Psychonomic Bulletin & Review*, *12*, 880-885.
- Fiser, J. z., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499-504.
- Fiser, J. z., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 458-467.
- Fiske, S., & Sanocki, T. (2010). Memory and attentional guidance in contextual cueing. *Journal of Vision*, *10*, 1302a.
- Geyer, T., Zehetleitner, M., & Müller, H.J. (2010). Contextual cueing of pop-out visual search: when context guides the deployment of attention. *Journal of Vision*, *10*, 1-11.
- Geng, J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*, 1252-1268.
- Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology*, *46*, 414-466.
- Goujon, A. (2011). Categorical implicit learning in real-world scenes: evidence from contextual cuing. *Quarterly Journal of Experimental Psychology*, *64*, 920-941.
- Goujon, A., Didierjean, A., & Marmèche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition*, *15*, 257-275.
- Goujon, A., Didierjean, A., & Marmèche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 50-71.
- Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision*, *9*(1):32, 1-8.
- Irwin, D. E., Andrews, R. V. (1996). *Integration and accumulation of information across saccadic eye movements*. Cambridge, MA, US: The MIT Press.
- Jiang, Y., & Song, J.-H. (2005). Hyperspecificity in visual implicit learning: Learning of spatial layout is contingent on item identity. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1439-1448.

- Jiang, Y., & Wagner, L. C. (2004). What is learnt in spatial contextual cuing—Configuration or individual locations? *Perception & Psychophysics*, *66*, 454–463.
- Jungé, J. A., Scholl, B. J., & Chun, M. M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition*, *15*, 1-11.
- Kunar, M. A., Flusberg, S. J., Horowitz, T. S., Wolfe, J.M. (2007). Does contextual cuing guide the deployment of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 816-828.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.
- Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, *109*, 376-400.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176-210.
- Oliva, A. & Torralba, A. (2007). The Role of Context in Object Recognition. *Trends in Cognitive Sciences*, *11*, 520-527.
- Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1299-1313.
- Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, *63*, 1239-1249.
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*, 17-42.
- Rogers, T. T. & McClelland, J. L. (2008). Precis of “Semantic Cognition, a Parallel Distributed Processing Approach”. *Behavioral and Brain Sciences*, *31*, 689-714.
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7*, 1-22.
- Smyth, A. & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition*, *36*, 403-415.
- Turk-Browne, N. B., Isola, P., Scholl, B., & Treat, T. (2008). Multidimensional Visual Statistical Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 399-407
- Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The Automaticity of Visual Statistical Learning. *Journal of Experimental Psychology: General*, *134*, 552-564.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience*, *30*, 11177-11187.
- Wolfe, J. M. (1998). Visual search. In h. Pashler (Ed.), *Attention* (pp. 13-73). London: University College London Press.

Accepted for publication: February 23, 2012