Implicit learning increases preference for predictive visual display

Hirokazu Ogawa · Katsumi Watanabe

Abstract We investigated whether implicit learning in a visual search task would influence preferences for visual stimuli. Participants performed a contextual cueing task in which they searched for visual targets, the locations of which were either predicted or not predicted by the positioning of distractors. The speed with which participants located the targets increased across trials more rapidly for predictive displays than for non-predictive displays, consistent with contextual cueing. Participants were subsequently asked to rate the "goodness" of visual displays. The rating results showed that they preferred predictive displays to both non-predictive and novel displays. The participants did not recognize predictive displays any more frequently than they did non-predictive or novel displays. These results suggest that contextual cueing occurred implicitly and that the implicit learning of visual layouts promotes a preference for visual layouts that are predictive of target location.

Keywords Visual preference · Contextual cueing · Implicit learning

H. Ogawa · K. Watanabe
Research Center for Advanced Science and Technology,
The University of Tokyo,
Tokyo, Japan

H. Ogawa (✉)
Department of Integrated Psychological Sciences,
Kwansei Gakuin University,
1-155 Uegahara, Nishinomiya-shi,
Hyogo 662-8501, Japan
e-mail: hirokazu.ogawa@kwansei.ac.jp

Introduction

In everyday life, continuous changes in the environment present the visual system with a vast amount of information. However, the visual system has severely limited resources, and it is therefore essential that information most relevant to current behavioral goals is selectively targeted and undergoes further processing. The visual system has two mechanisms of prioritization: visual attention and visual preference. Studies have identified parallels and interactions between these mechanisms and shown that they modulate one another. For example, Raymond, Fenske, and Tavassoli (2003) had participants perform a visual search task and then asked them to make affective evaluations of the targets and distractors that had just been presented and of novel stimuli. The distractors that participants had been asked to ignore during the search task were rated more negatively than targets or novel stimuli (see also Fenske, Raymond, & Kunar, 2004; Raymond, Fenske, & Westoby, 2005; for a review, see Fenske & Raymond, 2006).

The prioritization of information for processing by the visual system is regulated by learning and past experience. Events and changes in object properties do not occur in a random manner. Many studies of statistical learning have demonstrated that the human visual system extracts regularities and statistical structures inherent to a complex visual environment, which are learned implicitly to later provide contextual information that enhances visual performance (Chun & Jiang, 1998; Fiser & Aslin, 2001, 2002, 2005; Turk-Browne & Scholl, 2009). For example, the regular prediction of a target location by the spatial configuration of a display can be learned implicitly and visual search performance subsequently facilitated (contextual cueing; Chun & Jiang, 1998). In a typical contextual

cueing experiment, participants perform hundreds of visual search task trials in a manner that can be considered inefficient, because they require a serial deployment of attention. Targets are embedded either in configurations of distractors that are repeated (predictive layouts) across half of all trials or in novel layouts. Search reaction time (RT) is typically faster for targets in the predictive layouts than in the random layouts, and this difference increases as the task progresses. The performance benefit associated with the predictive spatial layouts is known as the "contextual cueing effect" and is considered to involve mechanisms that implicitly utilize the predictability of information in a visual scene to make attentional processes more efficient (for a review, see Chun, 2000).

In addition to an influence on visual attention, predictability may also modulate the extent to which certain visual stimuli come to be preferred over others. Indeed, it has recently been shown that participants had a tendency to like stimuli with predictive value more than stimuli without it (Bayliss & Tipper, 2006). They used the gaze-cueing paradigm to examine if impressions of a face were influenced by the extent to which its direction of gaze was predictive of target position. Throughout the trials, there was a consistent association between a particular face and the extent to which its direction of gaze predicted target location. Some faces always directed their eyes to the target location (predictive-valid faces), whereas others gazed to the opposite side (predictive-invalid faces). It was demonstrated that participants tended to evaluate predictive-valid faces as more trustworthy than predictive-invalid faces. Further, it was reported that very few participants mentioned the face-cue contingencies. These results suggest that the predictive value of faces was learned during the task and modulated preference judgments. However, since they measured the effect of the gaze cueing training on the preferential judgment by comparing between predictive-valid and predictive-invalid faces, the relationship between predictability and preferences could simply reflect a learned association between individual faces and predictability, and the extent to which the validity of gaze cues was involved is unclear. Moreover, it remains to be clarified whether the effect is specific for faces.

As mentioned above, there are often inherent regularities to the properties of objects present in everyday visual scenes. Previous research suggests that the predictable nature of these properties affects preferences for objects in such scenes. In the present study, we examined whether having predictable properties increases a preference for entire visual scenes. We used the contextual cueing paradigm to promote the implicit learning of predictable properties within visual scenes. Participants were repeatedly presented with two types of displays. The first was

predictive, with the configuration of distractor items and target position both being fixed across trials. The second was non-predictive, with the configuration of distractor items being fixed but the target position different on each trial. It was previously demonstrated that a contextual cueing effect is not produced when the configuration of a display is repeated but not predictive (Chun & Jiang, 1998, Experiment 3). By using the contextual cueing paradigm, we tried to dissociate the effect of predictability from that of repetitive presentation on preferences. The participants performed a visual search task, after which they were asked to evaluate the "goodness" of the displays that had been presented. If implicitly learned knowledge of predictability promotes a preference for visual displays, then predictive displays would be evaluated more favorably than non-predictive and novel displays. Alternatively, a preference for both predictive and non-predictive displays over novel displays would suggest the effects of mere exposure (Zajonc, 1968) rather than implicit learning.

Experiment 1

Method

Participants Twenty-three healthy young adults (aged 19–27 years old) with normal or corrected-to-normal visual acuity participated in the experiment after providing informed consent. The participants received 1,000 yen per hour and were not aware that the real aim of the experiment was to evaluate influences on the preference for visual stimuli.

Apparatus and stimuli The experimental tasks were programmed using MATLAB with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and presented on a 21-inch (c.53.3-cm) CRT Monitor (100 Hz refresh rate) viewed from 60 cm. Responses were made with a standard keyboard. We presented visual displays that consisted of 12 white squares on a gray background, including 11 distractors and one target (Fig. 1). Each square was $1.0° \times 1.0°$ of visual angle in size and had a gap (0.12°) on one side. For target squares, the gap was positioned on either the left or the right side, whereas for distractors, it was either at the top or the bottom. Each of the 12 squares was located within a cell of a 16×16 virtual matrix that subtended $12.5° \times 12.5°$ of visual angle. A total of 32 display configurations were generated at the beginning of each experimental session. For each configuration, the locations of the 12 items within the virtual matrix were randomly selected. Half the configurations were presented as predictive displays, in which the locations of both target and

Fig. 1 Examples of displays presented in the learning phase of the visual search task
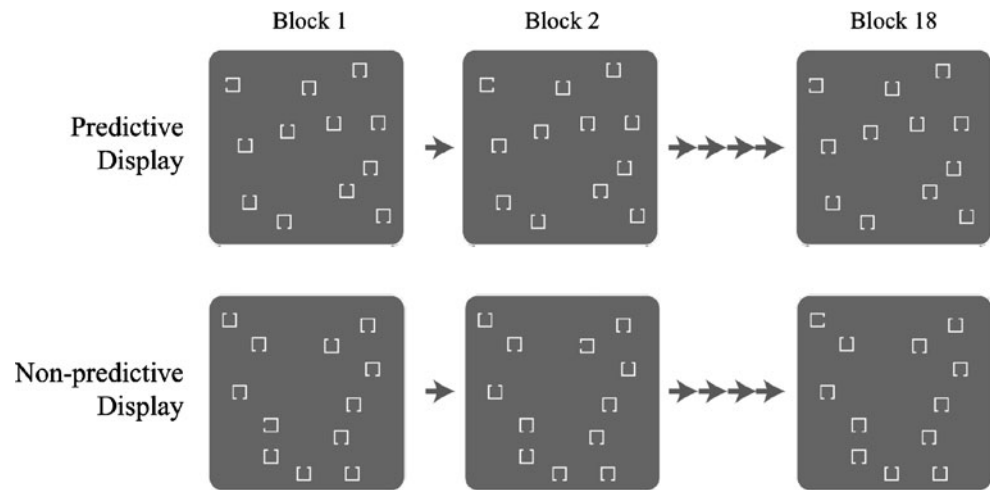
distractors remained constant across trials. The other half of the configurations were presented as non-predictive displays, in which the global configuration was fixed, but the target location was randomly assigned to one of the free matrix cells on a trial-by-trial basis.

Design and procedure Each experimental session consisted of 4 phases: learning (18 blocks of 32 trials), test (2 blocks of 48 trials), evaluation (2 blocks of 48 trials), and recognition (48 trials).

The learning and test phases involved a visual search task, trials of which began with a fixation display that was replaced after 750 ms with a search display. Participants were instructed to find the target among the distractors, and to indicate the orientation of its gap as quickly and accurately as possible. This was done by pressing the "F" and "J" keys for gaps to the left and right of a target, respectively. The display screen was cleared either upon a response being made or after a 5,000-ms response period had elapsed. Incorrect responses were signaled to participants by low-pitched tones. There was a 750-ms inter-trial interval in which only a blank screen was displayed. The identities of the targets and distractors were randomly selected for each trial. In the learning phase, each of the predictive and non-predictive display configurations was presented in a randomized order once per block of trials (and thus repeated 18 times). The test phase immediately followed after the learning phase. There were no additional instructions or announcements between the learning and test phases. As well as having fewer trials, the test phase differed from the learning phase in presenting 16 novel displays (the configurations for which were generated on a trial-by-trial basis) along with the predictive and non-predictive displays. The novel displays served as a baseline condition against which any contextual cueing effect arising from the learning phase could be measured.

Following the visual search task was the evaluation phase, in which participants were asked to evaluate the goodness of displays. Each of the predictive, non-predictive, and novel displays was presented once in each block of trials (the order of presentation was randomized), and the squares within the display were displayed without the gaps. The novel displays were newly generated for the evaluation phase and differed from those presented in the test phase. Displays were presented for 2,000 ms, after which they were cleared and an evaluation scale was presented in the center of the screen. Participants were requested to rate each display on a 5-point scale from 1 (very bad) to 5 (very good) by pressing the corresponding number key. The participants were free to use their own criteria for judging goodness, but when the task was introduced, unity, sparseness, global form of a configuration, and attractiveness were used as examples. No time pressure was exerted.

In the final phase of the experiment, the recognition phase, participants were informed that they would be presented with visual displays already seen in earlier phases of the experiment. The participants were instructed to indicate if they recognized the displays (by pressing one of two keys). There were 16 predictive and 16 non-predictive displays, and 16 novel displays newly generated for the recognition task; these displays were presented in a random order, and each display remained on the screen until a response was made. There was no time pressure on making responses, and feedback was not provided.

Results and discussion

Figure 2 shows the mean search RT for correct responses to predictive and non-predictive displays in each block of the learning phase. Error rates were very low (<1%), and statistical comparison between conditions was not made for these. For the purposes of analysis, the RT data were
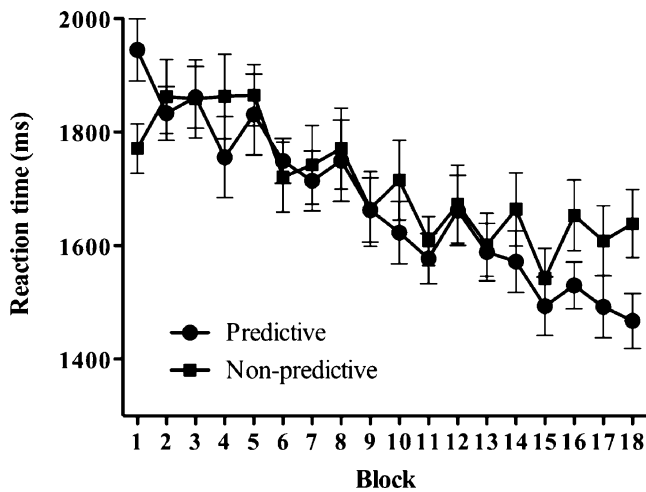
Fig. 2 Mean (±SEM) reaction times for the learning phase of Experiment 1

collapsed into six epochs of three consecutive blocks of trials (epoch 1 = blocks 1–3, epoch 2 = blocks 4–6, etc.). Repeated-measures analysis of variance (ANOVA) with display type (predictive and non-predictive) and epoch (1–6) factors was used to analyze the RT data. There was a main effect of epoch [$F(5, 110) = 26.62$, $p < .001$] and an interaction between display type and epoch [$F(5, 110) = 2.58$, $p < .05$]. A trend for a main effect of display type was also found [$F(1,22) = 3.03$, $p = .096$]. The results show that the performance of participants improved with time spent during the visual search task. Importantly, the level of improvement was greater for predictive displays than for non-predictive displays, suggesting that the participants learned the associations between target location and predictive display configuration.

Figure 3a shows the mean RT for each of the three display types in the test phase of Experiment 1. The ANOVA revealed a significant effect of display type [$F(2, 44) = 5.19$,

$p < .01$], with post hoc Tukey's HSD tests indicating that participants gave faster responses for predictive displays than for both non-predictive and novel displays ($ps < .05$); the RT did not differ between non-predictive and novel displays ($p = .98$). The test phase results of Experiment 1 were thus consistent with the development of the contextual cueing effect.

Figure 3b shows the mean rating of goodness provided by participants for each display type in the evaluation phase of Experiment 1. The ANOVA revealed a significant effect of display type on ratings [$F(2, 44) = 6.66$, $p < .005$], with post hoc tests indicating that predictive displays were rated more highly than both non-predictive ($p < .05$) and novel displays ($p < .01$); the ratings did not differ between non-predictive and novel displays ($p = .75$).

The mean proportion of trials on which participants indicated that they recognized a display is 59.0 (SD 10.8), 53.0 (SD 15.8) and 49.5% (SD 16.0%) (predictive, non-predictive and novel displays, respectively). Although the recognition rate for the predictive display seemed to be higher than non-predictive and novel displays, the main effect of display type was not significant [$F(2, 44) = 2.24$, $p = .11$], suggesting that predictability did not facilitate explicit recognition of previously presented visual displays. We suspect that this apparent increase in recognition rate may be attributable to the preceding evaluation phase wherein the participants scrutinized the predictive displays to evaluate their goodness.

Furthermore, to clarify whether preference was linked to the recognizability of the display, we compared the preference rates of the recognized predictive displays with those of the non-recognized predictive displays. The difference in the preference rating between the recognized and non-recognized displays was small (mean 3.04, SD 0.55 and mean 2.99, SD 0.61, respectively) and insignificant [$t(22) = 0.30$, $p = .77$]. This finding suggests that the

Fig. 3 a Mean (+SEM) reaction times for the test phase of Experiment 1. b Mean (+ SEM) ratings of goodness in Experiment 1. Ratings were from 1 (very bad) to 5 (very good)
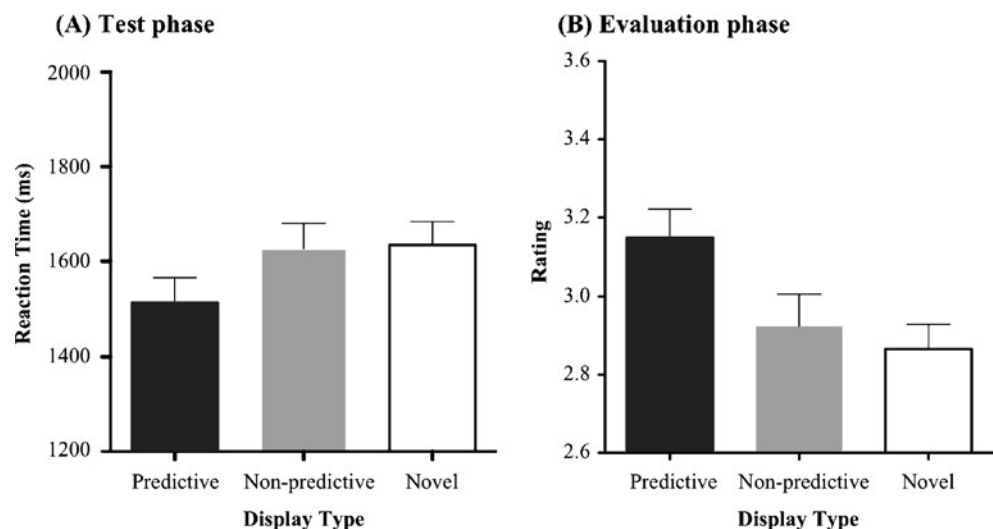
increased preference was not due to the recognition of the predictive displays.

The results of the test phase of Experiment 1 demonstrate the development of the implicit contextual cueing effect (e.g., Chun & Jiang, 1998). Most importantly, we found that participants evaluated predictive displays more highly than both non-predictive and novel displays. The development of preference for predictive displays cannot be accounted for by repetition alone, as non-predictive displays were not evaluated any more highly than novel displays. Accordingly, our results can be interpreted as evidence that implicitly learned knowledge of predictability influences preferences for visual stimuli.

However, there is an alternative mechanism by which our participants could have developed a preference for predictive displays. On the whole, visual searches were easier to complete with predictive displays than non-predictive or novel displays (an effect of contextual cueing). It is possible that positive emotions associated with the relative ease with which predictive displays were searched led to them to be evaluated more highly than other types of displays. This possibility was investigated in Experiment 2.

Experiment 2

The experimental procedures of Experiment 2 were similar to those of Experiment 1. An important exception was that the gap associated with target and distractor squares in the learning phase was larger in non-predictive displays than in both the predictive and novel displays. This made it easier for participants to find the target in non-predictive displays in Experiment 2 than in Experiment 1. If the preference for predictive displays found in Experiment 1 was because they were easier to search, then evaluations of predictive and non-predictive displays could be expected to be more similar in Experiment 2 than in Experiment 1.

Method

Participants Twenty-two healthy young adults (20–28 years old) with normal or corrected-to-normal visual acuity participated after providing informed consent. None of them participated in Experiment 1.

Apparatus and stimuli The size of the gaps associated with target and distractor squares in the non-predictive displays of the learning phase was 0.2° of visual angle (rather than the 0.12° used in Experiment 1). However, a gap size of 0.12° was used for all three display types in the test phase, enabling the development of a contextual cueing effect to be reliably assessed.

Design and procedure The order of experimental phases differed from Experiment 1, with the learning phase immediately followed by the evaluation phase (rather than the test phase). This was done so that evaluations were made before exposing participants (in the test phase) to non-predictive displays with a similar gap size to, and thus no longer easier to search than, predictive displays. Experiment 2 did not include a recognition phase, as we considered the results of Experiment 1, and their consistency with previous research, to be sufficient evidence for the implicit nature of contextual learning associated with our procedures.

Results and discussion

Figure 4 shows the mean search RT for correct responses to predictive and non-predictive displays in each block of the learning phase. As in Experiment 1, these data were collapsed into six epochs before being analyzed. The ANOVA revealed main effects for both display type [$F(1, 21) = 25.56$, $p < .001$] and epoch [$F(5, 105) = 27.02$, $p < .001$], and an interaction between these factors [$F(5, 105) = 3.15$, $p < .05$]. The main effect for display type is the most important of these results, as it verifies that RTs for non-predictive displays were faster than for predictive displays, indicating that searches of non-predictive displays were easier to complete.

Figure 5a shows the mean RT for each of the three display types in the test phase of Experiment 2. The ANOVA revealed a significant effect of display type [$F(2, 42) = 7.25$, $p < .005$], with post hoc tests indicating that participants gave faster responses for predictive displays than for both non-predictive ($p < .05$) and novel displays ($p < .01$); RT did not differ between non-predictive and novel displays ($p = .60$). As in Experiment 1, these results are
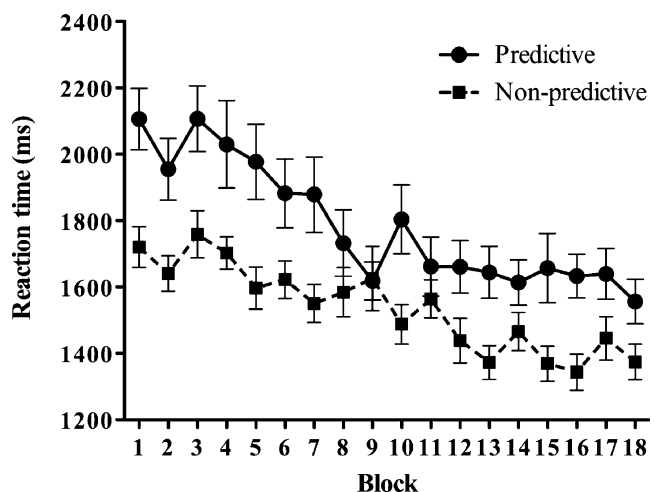
Fig. 4 Mean (±SEM) reaction times for the learning phase of Experiment 2

consistent with the development of the contextual cueing effect. They also suggest that the relative ease with which non-predictive displays were searched in the learning phase did not transfer to the test phase (in which the size of the gap associated with target and distractor squares was equal across predictive and non-predictive displays).

Figure 5b shows the mean rating of goodness provided by participants for each display type in the evaluation phase of Experiment 2. The ANOVA revealed an effect of display type on ratings [$F(2, 42) = 18.67$, $p < .001$], with Tukey's HSD tests indicating that predictive displays were rated more highly than both non-predictive ($p < .001$) and novel displays ($p < .001$). The difference between non-predictive and novel displays was not significant ($p = .11$).

In Experiment 2, we clarified whether the preference for predictive displays over non-predictive displays found in Experiment 1 was because they were easier to search (due to the contextual cueing effect) or because of the implicitly learned knowledge of predictability. Despite being easier to search, we found that non-predictive displays were evaluated less preferably than predictive displays. This result suggests that the ease with which visual displays could be searched did not influence how they were evaluated and supports a relationship between implicitly learned knowledge of predictability and preferences for visual stimuli.

General discussion

It has been demonstrated that repeated exposure to an initially unfamiliar stimulus enhances the extent to which it is subsequently favorably judged. This is known as the "mere exposure effect" (Zajonc, 1968), and is thought to be a result of enhanced processing fluency (Bornstein & D'Agostino 1994; Jacoby, Kelley, & Dywan, 1989; Whittlesea, 1993). In the present study, the participants

were repeatedly exposed to predictive and non-predictive displays to the same extent. Despite this, non-predictive displays were evaluated less highly than predictive displays, making it unlikely that the development of a preference for predictive displays was solely due to repetition. Thus, the present results suggest that the preference for predictive visual displays developed not because they were easier to search or were repeatedly presented, but because of the implicitly learned knowledge of predictability.

What is the mechanism by which implicit knowledge of predictability facilitates a preference for visual stimuli? A substantial body of research in both animals and humans has demonstrated an important effect of reward signals or reinforcement on learning (O'Doherty, 2004; Schultz, 2006), including perceptual learning (Seitz, Kim, & Watanabe, 2009; Seitz & Watanabe, 2005). Contextual cueing can be considered as a type of associative learning (Chun & Jiang, 1998), and it is therefore likely that reinforcement processes are involved. Contextual cueing facilitates target detection. The successful performance of a task serves as a reward (Karni & Bertini, 1997; Herzog & Fahle, 1999), and the detection of a target is thus likely to generate reward signals that feedback to reinforce contextual cueing processes. We propose that these reward signals may also promote a preference for the visual stimuli with which they are associated (e.g., the predictive displays of the present study). The extent to which contextual cueing and other types of implicit learning share reinforcement mechanisms is unclear. Nevertheless, clarifying the involvement of reward signals in contextual cueing may help in understanding how reinforcement mechanisms promote implicit learning.

In contrast to the large body of literature on the mere exposure effect, our data did not show increased preference for the non-predictive displays even though these displays were repeated as often as predictive displays. Moreover, no difference was observed between novel and non-predictive

Fig. 5 a Mean (+SEM) reaction times for the test phase of Experiment 2. b Mean (+SEM) ratings of goodness in Experiment 2. Ratings were from 1 (very bad) to 5 (very good)
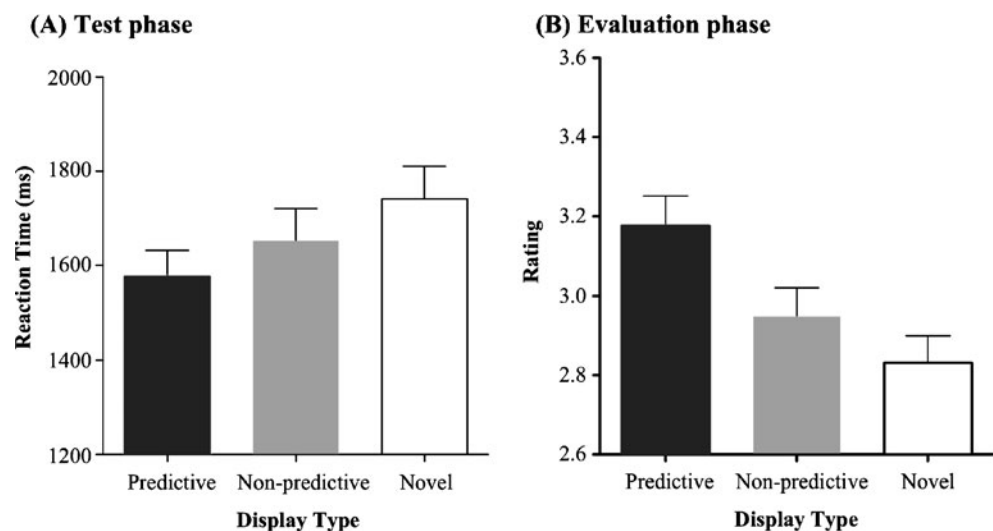
displays even when the displays were associated with easier search (Experiment 2). One possible explanation is that it may be due to erroneous attentional guidance in the non-predictive displays. That is, in non-predictive displays, the configuration was consistent but the target location varied for each presentation. Thus, learning of the repeated configuration may guide attention to a location where the target did not exist. This failure of guidance may produce "punishment signals" and might have a negative impact on the preference for the display. This negative impact might cancel out the mere exposure effect or the effect of search efficiency on preference for the non-predictive displays. The preference for a visual display could be influenced by numerous factors. An important goal for future research is to elucidate the mechanisms by which such factors interact with each other to modulate visual preference.

There may also be other mechanisms whereby contextual cueing procedures generate affective responses. Recent studies have described how "distractor devaluation" can affect preferences for visual stimuli (Fenske & Raymond, 2006). In the present study, any devaluation effect would appear to have been inconsequential in comparison to the contextual cueing effect generated through repeated exposure to predictive visual scenes. Nevertheless, it has been shown that the attentional processes underpinning contextual cueing include both facilitation at target locations and inhibition at distractor locations (Ogawa, Takeda, & Kumada, 2007). This could mean that targets with a predictable location make a greater contribution to affective responses than the distractors by which target location can be predicted. The extent to which any such process occurs and its role in the development of preferences for predictable visual stimuli are topics for future research.

Our demonstration that predictability promotes a preference for visual scenes has important implications for understanding how visual information in complex real-world scenes comes to be prioritized. In a visual world with rich depth and perspective, there are numerous possible viewpoints and locations from which a search for behaviorally relevant objects can be performed. A preference for particular visual stimuli can help an observer identify a vantage point that enables them to perform an efficient visual search. In other words, perspectives that are preferred because of their predictability are more likely to be chosen for searching, with contextual information subsequently guiding the allocation of attention within the scene.

References

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science, 17*, 514–520.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition, 12*, 103–128.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*, 170–178.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71.

Fenske, M. J., & Raymond, J. E. (2006). Affective influences of selective attention. *Current Directions in Psychological Science, 15*, 312–316.

Fenske, M. J., Raymond, J. E., & Kunar, M. A. (2004). The affective consequences of visual attention in preview search. *Psychonomic Bulletin & Review, 11*, 1055–1061.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499–504.

Fiser, J., & 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.

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General, 134*, 521–537.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In F. I. M. Craik & H. L. Roediger (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Lawrence Erlbaum.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: behavioral probes into adult cortical plasticity. *Current Opinion in Neurobiology, 7*, 530–535.

Herzog, M. H., & Fahle, M. (1999). Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vision Research, 39*, 4232–4243.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology, 14*, 769–776.

Ogawa, H., Takeda, Y., & Kumada, T. (2007). Probing attentional modulation of contextual cueing. *Visual Cognition, 15*, 276–289.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Raymond, J. E., Fenske, M. J., & Tavassoli, N. (2003). Selective attention determines emotional responses to novel visual stimuli. *Psychological Science, 14*, 537–542.

Raymond, J. E., Fenske, M. J., & Westoby, N. (2005). Emotional devaluation of distracting patterns and faces: A consequence of attentional inhibition during visual search? *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1404–1415.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology, 57*, 87–115.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron, 61*, 700–707.

Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Science, 9*, 329–334.

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 195–202.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*, 1235–1253.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, Monograph Supplement, 9*, 1–27.