LEARNING IDENTITY-TO-LOCATION BINDING WITH FACES

DOES IT MATTER WHO WAS WHERE? LEARNING IDENTITY-TO-LOCATION BINDING FROM FACES

By

MICHAEL JUSTIN WAN, B.Sc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree of Master of Science

Descriptive Note

McMaster University MASTER OF SCIENCE (2019) HAMILTON, ONTARIO (Department of Psychology, Neuroscience & Behaviour)

TITLE: Does it Matter Who was Where? Contextual Cueing with Faces AUTHOR: Michael Justin Wan, B.Sc (McMaster University) SUPERVISOR: Dr. Hongjin Sun

NUMBER OF PAGES: vii, 65

Abstract

People unconsciously learn spatial information about places they encounter frequently, leading them to search through familiar scenes faster than for unfamiliar scenes. We explored this phenomenon—the contextual cueing effect—in scenes containing images of different human faces. Participants searched through a series of scenes for a target among distractors, characterized as a letter *T* among letter *L*'s with each letter positioned on top of a face image (Experiment 1) or as a female face among male faces (Experiment 2). Experiment 1 showed that when the binding of identity and location was manipulated during learning, slightly greater (but not statistically significant) contextual cueing effects were found for repeated scenes with constant identity-to-location binding than those repeated scenes with constant spatial configurations but shuffled identity-to-location binding. Experiment 2 showed that if the binding of identity-to-location changed after the learning of a set of identity-to-location binding, small (but not statistically significant) costs of contextual cueing were found. The results suggest that in the contextual cueing paradigm, repeated identity-to-location binding might be learned but the learning of repeated spatial configurations alone account for a major portion of the learning.

Acknowledgments

I wish to begin by thanking my supervisor, Dr. Hongjin Sun, for his counsel and support throughout this project. His dedication and thought-provoking insights have been an invaluable source of inspiration to me.

I also thank my committee members: Dr. David Shore and Dr. Judith Shedden for their feedback and suggestions, further enriching the research process, from data collection to writing.

I thank Dr. Xiaoqing Gao and Dr. Mel Rutherford for allowing me to access their banks of visual stimuli. I am grateful for Dr. Chao Wang and Ms. Zheng Lei for instructing me in using and troubleshooting their experiment software, providing me with some of the experiment stimuli, and for their excellent counsel for research standards. I thank Laura Jin for her time spent on data collection. I thank my fellow lab mates, Nadia Wong, Jiali Song, Hanna Haponenko, and Ben Sclodnick for the many fruitful conversations and for motivation. I am grateful for the members of the Sun Lab, from whom I have further deepened my appreciation for science and learning.

Finally, I wish to thank my family: Jane Wang, Calvin Wan, and Peter Wan for their constant support and encouragement. I am also grateful to my friends for having faith in me, and for their motivation to continue persevering.

Table of Contents

List of Figures and Tables

List of abbreviations

CCE = Contextual cueing effect

CC = Contextual cueing

RT = Reaction time

ms = milliseconds

GENERAL INTRODUCTION

People repeatedly experience different scenes as part of their daily routines, where each scene has a unique spatial layout and a unique set of objects. For example, most individuals would consider a room with scattered items and belongings to be messy. The room's tenant, however, may view their room as an intricate arrangement of objects that should not be disturbed. If the room is left untouched, the tenant would find particular items much faster than those unfamiliar with the room. In controlled visual search tasks, viewers find a target among distractors faster in previously encountered search scenes, than in novel scenes (Chun & Jiang, 1998). This phenomenon—known as the contextual cueing effect (CCE)—is an implicit process (Chun & Jiang, 2003) that facilitates visual search for familiar scenes.

1. Contextual cueing effect: traditional paradigm—learning target locations in reference to their distractor spatial layout

1.1 The mechanisms and nature of contextual cueing

When viewing a familiar scene, how do people find what they are looking for so quickly? Chun and Jiang (1998) characterized contextual cueing (CC) as the learning of target-to-context associations. In their classical visual search paradigm, viewers search hundreds of scenes for a target '*T*' among rotated distractor '*L*'s and declare the target's pointing direction. While every scene in each block features its own distractor layout and target location, half of the scenes reappear once in every block, and the remaining scenes are only shown once. Over time, people gain an advantage for searching for targets within *repeated* (or '*old*') scenes over *non-repeated* (or '*new*') scenes, but most are unaware that any spatial learning has occurred. Following the experiment, participants were asked if they noticed any patterns or regularities in the scenes—most participants reported that they did not notice any repetition. In a follow-up recognition test, the participants correctly declared repeated scenes as repeated at chance level. Taken together, this implies that their learning was largely implicit (Chun & Jiang, 1998; see also Chun & Jiang, 1999).

1.2 Specificity of context-to-target location associations

The contextual cueing effect from a repeated scene's spatial layout depends specifically on the relation between the scenes' distractor layout and their targets' locations. If a '*repeated*' scene appears each time with an unfamiliar target location, viewers may lose their search advantage for that repeated scene. For example, in Chun & Jiang (1998, Experiment 3) and Makovski (2017, Experiment 2), repeated scenes' target positions were integrated into the overall object configuration—each time an '*old*' scene appeared, the target would occupy one of the distractor's positions from when the scene was previously seen. Therefore, each repeated scene's "shuffled" context—which consisted of its own set of target and distractor positions—cued a different position with each presentation. In both cases, this '*integrated*' configuration did not facilitate search, as participants failed to establish CC for repeated scenes (Chun & Jiang, 1998; Makovski, 2017). In Chun & Jiang's (1998) 6th experiment, the repeated scenes' initial target positions were separate from their respective distractor configurations. During the first phase of the experiment, each repeated scene cued two possible target positions rather than one. Towards the end of the experiment, the repeated scenes' targets replaced random distractors. Participants still established CC for repeated scenes during the first phase of the experiment, but not during the second phase (Chun & Jiang, 1998, Experiment 6). Taken together, this suggests that the spatial learning of a repeated scene depends on the association between its target and its distractor layout. If a scene's distractor layout constantly changes—such as when the target replaces a distractor at a random position—then the target-to-context association will change each time the scene appears.

1.3 Multiple target positions in repeated scenes

As noted above, Chun & Jiang (1998, Experiment 6 first phase) observed CC for repeated scenes with two possible target positions, but also reported weaker to non-significant CC for repeated scenes with three possible target positions. Thus, viewers still demonstrated spatial learning despite the fact that each repeated scene's distractor configuration was associated with more than one target position. However, these target positions did not occupy previous distractor positions, as in Chun & Jiang (1998, Experiment 3), and in Makovski (2017, Experiment 2). Goujon, Didierjean, and Poulet (2014, Experiment 2) tasked participants to search through a photo of a real-world scene for a lone letter *T* or *L*, and declare its identity. Half of the scenes were repeated, but each target location in each scene could appear in one of eight possible locations. Participants did not gain any advantage for searching repeated scenes over searching non-repeated scenes (Goujon, et al., 2014). The real-world scenes in Goujon et al.'s experiment were rich in visual information—which could have contributed to the lack of search facilitation—compared against the blank background in Chun & Jiang's (1998) experiments. Moreover, the scenes in Goujon et al. (2014, experiment 2) featured eight possible target locations, while the scenes in Chun & Jiang (1998, experiment 1) held only one target location each. Regardless, it may be possible to observe CC effects for 'dynamic' repeated scenes—repeated scenes with more than one possible target position—but it is difficult to establish and maintain. Moreover, the CCE weakens and becomes less likely to manifest as the number of possible target locations in a repeated scene increases.

1.4 Disruption of contextual cueing for repeated scenes

Individuals can lose CC for repeated scenes if their respective context-to-target associations are disrupted. Participants completed a classical *T* and *L* visual search task—search for the letter *T* and declare its pointing orientation upon discovery—except that all of the repeated scenes' targets were moved halfway through the experiment (Conci, Sun, & Müller, 2011). Each repeated scene's target was moved to a previous distractor position—while the remaining distractor layout was left intact—and remained in that position for each subsequent presentation until the end of the experiment. During the first half of the experiment, participants developed strong contextual cueing effects for the initial sets of repeated target locations and their respective distractor layouts. After each repeated scene's respective target switched positions with that of a distractor, participants lost their search advantage for repeated scenes, and the contextual cueing effect failed to recover, even after sufficient re-training (Conci et al., 2011). It is interesting that viewers were not able to regain any contextual cueing effects for repeated scenes following the change in target positions, despite only occurring once during the entire experiment. According to Conci et al. (2011), CC did not re-emerge because the repeated scenes' targets moved to previous distractor positions. Although the repeated scenes' targets were shuffled only once, their new positions changed their respective scenes' distractor layouts. Since the new repeated scenes' distractor layouts included positions formerly associated with targets, it is possible that this hindered search facilitation, even without any further changes in target positions. Makovski and Jiang (2010) found similar results by modulating target positions—as a repeated scene's target moves farther away from its original position, any previous contextual cueing effects for that scene continue to weaken until they are lost. It is important to note that the repeated scenes' target positions were moved to unoccupied positions, rather than previous distractor positions. While the distractor layout was left intact, the association between a scene's target and its respective context was disrupted until the scene no longer facilitated search (Makovski and Jiang, 2010). Therefore, even already-established contextual cueing effects can deteriorate or disappear if the repeated contexts are disrupted.

In summary, the establishment of location-based contextual cueing for a repeated scene depends on its repeated element. When a repeated scene's target and distractor spatial configuration are fixed, contextual cueing is strengthened (Chun & Jiang, 1998). When a repeated scene's target position changes each block, contextual cueing is worsened (even if the scene's distractor configuration is fixed). Contextual cueing is strengthened when a repeated scene's target always appears in its original position. Contextual cueing is worsened when a repeated scene's target appears sufficiently far away from its original position, in relation to the scene's fixed distractor configuration (Makovski & Jiang, 2010). Lastly, contextual cueing is strengthened when a repeated scene's target occupies a unique position, separate from the positions constituting the fixed distractor configuration. Contextual cueing is worsened when a repeated scene's target position replaces a random distractor position (Chun & Jiang, 1998; Conci et al., 2011, Makovski, 2017).

1.5 Switching target locations between repeated contexts

Despite the specificity in CC, there is some flexibility with the search benefit for repeated scenes. In Zellin, von Mühlenen, Müller, & Conci (2013), viewers first established CC with a set of repeated scenes in a learning phase—each scene had its own distractor layout and target position—then all of the repeated scenes' target positions were shuffled (the exchange phase), such that each repeated scene's target was positioned at a former target position from a different repeated scene (the return phase). During the exchange phase, the viewers' search benefits for repeated scenes only diminished slightly before continuing to improve, and remained constant as the repeated scenes' target positions reverted back to normal (the return phase). An important distinction here is the difference in target-to-context associations in the repeated scenes. Earlier paradigms such as Chun & Jiang (1998, experiments 3 and 6), Conci et al. (2011), and Goujon et al. (2014) involved repeated scenes' target positions replacing former distractors at some point during the experiment. By contrast, all of the repeated scenes in Zellin et al. (2013) used the same set of target positions and sets of distractor positions throughout the experiment. Therefore, the exchange phase's distractor layout and target pairings did not disrupt the CCE from the learning phase. This suggests that the repeated exposure to a set of distractor layouts and a set of target locations can facilitate search, provided that a particular repeated distractor layout is paired with any repeated target position (Zellin et al., 2013).

1.6 Attended vs. ignored contextual information and spatial learning

While contextual cueing for a scene requires repeated presentations of the same spatial layout of items, individuals will not manifest CC unless they attend to that spatial context. Jiang & Chun (2001) used a modified *T* and *L* visual search paradigm to investigate the effects of attention on CC. Viewers searched for a target *T* among distractor *L*'s, and declared the target's pointing orientation, but half of the letters were coloured red, and the remaining letters were coloured green. The viewers were informed that their target would always be one of the above two colours for the duration of the experiment. Thus, the letters that were of the same colour as the target were considered the "*attended*" context, and those that were of the opposite colour as the target were considered the "*ignored*" context. In Jiang & Chun's (2001) first experiment, each repeated scene's "*attended*" context was kept constant ("*attended-old*"). Although it emerged during the final blocks of the experiment, viewers did establish CC with the *attended-old* context. In contrast, viewers did not establish CC in Jiang & Chun's (2001) second experiment, where each repeated scene's "*ignored*" context was kept constant ("*ignored-old*"). Both types of repeated scenes included some variation in their overall spatial layout, but CC only manifested for attended-old scenes. Taken together, this suggests that viewers must attend to relevant spatial information within a repeated scene in order for search facilitation to occur.

1.7 Learners vs. Non-learners

Repeated exposure to familiar scenes promotes faster subsequent search, but not all individuals necessarily learn the repeated context to the same extent. For a given visual search paradigm, for the same amount of exposure, some viewers will successfully learn the repeated contexts (i.e. "learners"), while other viewers will fail to learn them (i.e. "non-learners"). To determine whether or not viewers learn repeated contexts, their reaction times for non-repeated scenes are compared against those for repeated scenes. Chun & Jiang (1998) compared reaction time differences between scene conditions from the later stages of the experiment (over the last three epochs) against zero milliseconds. Lleras & von Mühlenen (2004) used this calculation as an exclusion criterion to assess changes to learning between their learning phase and transfer phase. Only participants that established a positive contextual cueing were included in their transfer phase analysis. Other researchers have used this calculation to describe contextual cueing with different time frames within their experiments. Conci et al. (2011) and Zellin, Conci, von Mühlenen, and Müller (2013) have used the same calculation to assess contextual cueing at the block level. Olson, Chun, and Allison (2001), Kunar & Wolfe (2011), Conci & Müller (2012), Zellin et al. (2013), have used this calculation to assess contextual cueing at the epoch level. The distinction between "learners" and "non-learners" allows for a deeper insight into how likely individuals would successfully learn repeated contexts, and the extent to which learning facilitates future search (or the extent to which a lack of learning hinders future search).

2. Learning of repeated identity of search items, in the absence of repeated spatial layout

Repeated scenes continuously present complex spatial information, but they also carry visual information about the search items within the 'context.' This information—the target and distractors' identities—can possibly be learned and facilitate visual search, if repeated. Using abstract shapes, Chun & Jiang's (1999) search task involved finding the lone shape with vertical symmetry. Half of the scenes' target and distractor sets—instead of their target locations and distractor layouts—were repeated once each block, while the remaining scenes were presented only once. Therefore, the identity information from each repeated scene was predictive of the scene's target identity, not its location. The repeated identity information was a strong cue for target identity, as observers still established CC over time (Chun & Jiang, 1999). Endo & Takeda (2004, Experiment 2) also found CC in search scenes with free-drawn contours, when each repeated scene only repeated its own set of distractors (rather than maintaining the distractor's spatial arrangement). However, the repeated item identities in each repeated scene facilitated the search of the target location, rather than its identity. It is interesting to note that contextual cueing still occurred in both instances—in both studies, viewers could not make any reliable associations between the repeated scenes' target locations and their respective distractor configurations, due to their non-repeated arrangements. Instead, contextual cueing emerged when repeated distractor identities were predictive of target identities (Chun & Jiang, 1999), and when repeated distractor identities were predictive of target locations (Endo & Takeda, 2004).

Fixed target and distractor identity information also improved search efficiency for repeated visual search for real-world objects. Hout & Goldinger's (2010, experiments 1C and 2C) visual search paradigm consisted of a set of repeated scenes appearing in random sequence for each block. Participants searched through each scene either to declare whether or not a target was present, or which out of a set of possible targets was present. In both experiments, one of the conditions held the item identities for any particular scene constant (except the target, which was different for every trial). The items in each scene for this condition had a random spatial layout. Still, viewers gradually improved their search times for these *identity*-repeated scenes. Due to the low probability of recurring local or global spatial layouts over the full sequence of trials, Hout & Goldinger (2010) attributed the improvement in search times to the learned object identities. Although the nature of learning identity information is distinct from conventional CC (learning spatial information), viewers can gain an advantage for searching scenes with repeated object identities.

3. Learning of repeated identity-to-location binding

As discussed above, repeated object identities and repeated object locations both contribute to search facilitation. The repetition of object identities and object locations together introduces a third context: identity-to-location binding. A repeated scene's identity-to-location binding is considered repeated (constant), when both location information and identity information are held constant, such that each specific distractor always occurs at a specific location. To investigate the learning of identity-to-location binding as a repeated context, single-phase and dual-phase paradigms can be used.

3.1 Single-phase vs. dual-phase visual search paradigms

3.1.1 Single-phase paradigms

Experiments with single-phase paradigms allow for investigating possible search facilitation for scenes without repeated identity-to-location binding. Repeated scenes would hold location information constant, and also hold identity information constant for the entire set of search items. The correspondence between a given item's identity and its given location, however, would constantly shuffle across blocks (i.e. "*repeated-unbound*"). If learning binding information is required for search facilitation, participants would not be faster in searching through repeated scenes than for non-repeated scenes. In this case, the lack of binding between the two repeated contexts would interfere with participants' learning of the repeated location information. Conversely, if learning binding information is not required for search facilitation, participants would consistently search through repeated scenes faster than non-repeated scenes, as location information would be sufficient to establish CC.

3.1.2 Dual-phase paradigms

Repeated scenes in the dual-phase paradigms hold identity-to-location binding information constant during the initial, "learning" phase. Following the learning phase, the distractors in each repeated scene would be shuffled, and search performance would be examined at the beginning of the transfer, "testing" phase. In contrast to the single-phase paradigm, the experiments with dual-phase paradigms investigate viewers' likelihood of maintaining CC, after establishing it. The fixed identity-to-location binding during the learning phase allows participants to establish CC, regardless of whether or not learning binding information is required. The second phase provides a new set of binding between repeated location information and repeated identity information. If participants learn the repeated binding information during the learning phase, then they would be slower to search through the repeated scenes during the testing phase. Therefore, the shuffling of binding between the two phases would successfully interfere with search facilitation. Conversely, if participants do not learn the repeated binding information during the learning phase, then their search performances for repeated scenes during the testing phase would be unaffected. In this case, the shuffling binding between the two phases would fail to interfere with search facilitation.

3.2 Identity-to-location binding and search facilitation with abstract contours

Endo & Takeda (2004, Experiment 1) investigated search facilitation using a dual-phase, within-subjects paradigm, and used abstract contours for their search arrays. Each search array displayed 12 contours, drawn from a pool of 96 contours. During the learning phase, the repeated search scenes bound and held constant their respective distractor configurations and distractor identities (i.e. the "*combined repetition*" condition) between blocks. During the testing phase, a set of scenes from the combined repetition condition were altered, introducing a new "*configuration repetition*" condition. For configuration repetition scenes, the distractor configurations were repeated between blocks, but their respective sets of distractor identities were changed between blocks. Thus, for a given configuration repetition scene, its set of distractor identities was a set of 11 items drawn randomly from the stimulus pool of 96 items, each block. Contextual cueing remained for scenes in the configuration repetition condition, which suggested that: 1) repeated location information from the learning phase was sufficient for search facilitation to occur during the testing phase; and 2) identity information was not learned during the learning phase (Endo & Takeda, 2004).

Endo & Takeda (2004, Experiment 3) further explored the combined repetition and the configuration repetition conditions, using a single-phase, within-subjects design. For scenes in the combined repetition condition, their respective sets of distractor identities remained unchanged between blocks. For scenes in the configuration repetition condition, their respective sets of distractor identities were changed each block, with randomly drawn sets of 11 items from the stimulus pool of 96 items. Search facilitation occurred for scenes in the configuration repetition condition, suggesting that inconsistent distractor identities do not hinder learning. Search facilitation was stronger for scenes in the combined repetition condition, where the binding information was repeated. For abstract contours, this suggests that location information is sufficient for contextual cueing, but that repeated binding information can strengthen contextual cueing. (Endo & Takeda, 2004, Experiment 3).

3.3 Identity-to-location binding and search facilitation with real-world objects

Identity-to-location binding has also been investigated for real-world objects. Makovski's (2016, Experiment 1) featured a single-phase, within-subjects design, where search performances were compared for scenes in the "*all-repeat*" condition, where distractor configurations and corresponding distractor identities were bound and held constant between blocks; and scenes in the "*location-repeat*" condition, where the distractor configurations—but not the corresponding distractor identities—were repeated between blocks. Both Makovski's (2016) and Endo & Takeda's (2004) had similar set sizes for their search arrays (15 vs. 12 items, respectively). But Makovski's (2016) stimulus pool consisted of 2400 items, where each participant would view a random sub-pool of 350 items—much larger than Endo & Takeda's (2004) stimulus pool of 96 items. The larger stimulus pool in Makovski (2016) increased the likelihood for a location-repeat scene to feature novel distractor identities each block than a configuration repetition scene in Endo & Takeda (2004). Therefore, the location-repeat condition in Makovski (2016) offered a stronger test for learning repeated location information.

Contextual cueing was present for scenes in the all-repeat condition, but was absent for scenes in the location-repeat condition. Contrary to Endo & Takeda's (2004) observations for abstract contours, search facilitation for real-world objects occurred only in scenes where binding information was repeated (Makovski, 2016).

The requirement of binding for search facilitation was further investigated in Makovski (2017, Experiment 2). Using a single-phase, within-subjects design, search performance for scenes in an "*unbound*" condition were compared against non-repeated, "*new*" scenes. Scenes in the "*unbound*" condition repeated both their respective distractor configurations and distractor identities—unlike scenes in Makovski's (2016) location repeat condition, which only preserved the distractor configuration—but the binding between the two repeated contexts shuffled between blocks. Contextual cueing was absent for the unbound scenes, which further supported that learning repeated binding information was necessary to facilitate search for real-world objects (Makovski, 2017).

3.4 Stimulus properties: viewer expertise, uniqueness, semantic information, and variety

Contextual cueing studies in the literature have established learning using several different stimulus types. Some of these types include rotated letters (e.g. Chun & Jiang, 1998; Jiang & Chun, 2001; Makovski & Jiang, 2010; Conci et al., 2011; Zellin et al., 2013), abstract contours (Endo & Takeda, 2004), words (Goujon et al., 2009), and images of real-world objects as search stimuli (e.g. Makovski, 2016; Makovski, 2017). One can distinguish between different stimulus types using three categories: viewer expertise—the relative level of exposure and experience people are likely to have with a given stimulus type (high or low); uniqueness—the range of visual differences between images within the same semantic category (high or low); semantic information—the relative range of extractable characteristics or features from a given stimulus type (high or low); and variety—the size of the stimulus set, defined by the number of unique images that can appear in the experiment (high or low). For example, most individuals only encounter capital *T*'s and capital *L*'s in their upright position, rather than in their rotated or flipped positions. Therefore, it is reasonable to argue that individuals have low viewer expertise for rotated letters. The rotated letters used in most contextual cueing studies appear in a limited number of possible angles (e.g. Chun & Jiang, 1998; Jiang & Chun, 2001; Makovski & Jiang, 2010; Conci et al., 2011; Zellin et al., 2013), such that these letters have low uniqueness. Since individuals likely have viewed rotated capital *T*'s and capital *L*'s less often (compared against their upright counterparts), the rotated letters have a low level of semantic information. In the literature, the rotated letters have had varying relative levels of variety. In Chun & Jiang (1998), each of the letters was coloured one of: blue, green, yellow, or red. Jiang & Chun's (2001) letters were coloured either green or red, and Makovski & Jiang's (2010) letters were all coloured white. Of these three studies, Chun & Jiang's (1998) letters had the most variety, while Makovski & Jiang's (2010) letters had the least variety. Using these four defining criteria, novel contours have low viewer expertise, low uniqueness, low semantic information, and high variety. Similarly, words (for one's native language) have high viewer expertise, high uniqueness, high viewer expertise, and high variety. Images of real-world objects have high viewer expertise, high uniqueness, high semantic information, and high variety.

3.5 Contextual cueing established from faces

Most visual search studies investigated learning for spatial contexts, using rotated *T*'s and *L*'s for search stimuli. Visual search studies investigated learning for identity contexts used a wider variety of stimuli, including abstract contours (Endo & Takeda, 2004) and real-world images (Makovski 2016; Makovski 2017). Endo & Takeda (2004) observed weak learning for identity information with abstract contours, while Makovski (2016; 2017) observed strong learning for identity information with real-world images. It can be argued that the strength of learning for item identity depends on the viewer expertise for and semantic information held by a given stimulus type. Makovski's (2016) real-world image stimuli were rich in viewer expertise and in semantic information. But it was not apparent as to which property—the viewer expertise, the semantic information, or both—contributed to learning for repeated identity information.

Zheng et al. (2019) used faces as search items to investigate implicit learning for identity information, where the search items had high viewer expertise but carried little semantic information (unlike real-world objects, belonging to different categories, e.g., backpacks, computers, bottles, etc., as in Makovski (2016) and Makovski (2017)). While participants had higher viewer expertise for upright faces, they had lower viewer expertise for inverted faces. Individuals fail to recognize familiar faces from novel foils more often when the faces are inverted than when the faces are upright (Hochberg & Galper, 1967). Moreover, inversion hinders recognition for faces more than recognition for other objects (Yin, 1969). This suggests that viewers are more likely to learn repeated identity information from upright faces than from inverted faces. When investigating identity information learning, inverted faces can serve as a control for upright faces, because viewers will likely fail to learn repeated contexts from inverted faces, but will likely succeed in learning repeated contexts from upright faces. The comparison of search performances between inverted and upright faces would allow researchers to isolate the effect of viewer expertise on learning.

Zheng et al. (2019, experiment 1A) used a dual-phase design to investigate the potential for learning repeated binding contexts. In previous studies (Endo & Takeda, 2004; Makovski, 2016; 2017) the search items in each scene were drawn from a larger pool of items used in the experiment (i.e. different scenes involved different sets of items). Zheng et al. (2019) took a different approach by using the same set of items in each scene, which consisted of twelve upright, different (i.e. unique) faces. The restrictive stimulus pool in Zheng et al. (2019) offered a stronger test for binding learning, as participants were exposed to the same object identities more often than in Endo & Takeda (2004) and in Makovski (2016; 2017).

In Zheng et al.'s (2019) first experiment, eleven of the scene's search items were marked as distractors, with a rotated letter '*L*' appearing above the faces' foreheads. The target item was marked with a rotated letter '*T*' appearing above the face's forehead. Participants searched for the target face and declared the pointing direction of the letter '*T*'s' stem. Half of the trials in each block repeated once every block throughout both phases, and the remaining trials only appeared once over the duration of the experiment. Each repeated scene held its distractor configuration—bound with its set of distractor identities (identity-to-location binding)—and its target location constant. For the testing phase, the faces in each repeated scene were shuffled, but maintained the same respective spatial layout. The search performances for scenes with the upright, different faces stimulus set were compared against three other between-subjects conditions: upright, same faces; inverted, different faces; and inverted, same faces. Results showed that during the learning phase, viewers in all conditions established a stable contextual cueing effect. During the first block of the testing phase, search performance for upright, different faces deteriorated for the repeated scenes, eliminating the CCE. The CCE then re-emerged and remained until the end of the testing phase. For the remaining three conditions, search performance was unaffected between the end of the learning phase and the beginning of the testing phase. This suggested that only viewers in the upright, different faces condition learned the repeated binding information, and that this learned repeated binding information interfered with the viewers' search performances for repeated scenes in the testing phase (Zheng et al., 2019, Experiment 1A). This conclusion was further supported in Zheng et al.'s (2019) second experiment, which was conducted using the same design as the first experiment, but with a different search task: searching for the lone male face out of the distractor female faces. Binding learning occurred during the learning phase, as the CCE for scenes in the different-upright faces condition was eliminated at the beginning of the testing phase (Zheng et al., 2019, Experiment 1B).

A third experiment by Zheng et al. (2019) examined the effects on spatial learning for faces, with repeated location information, but inconsistently paired with repeated identity information. The experiment used a single-phase, between-subjects paradigm, and compared contextual cueing effects for two types repeated scenes: "*unbound*" and "*bound*" (Zheng et al., 2019, Experiment 2). Search arrays consisted of eleven distractor female faces, and one target male face. Half of the scenes featured identical distractor face, and the other half of the scenes featured different distractor faces. For scenes in the unbound condition, the spatial layouts were held constant between blocks, but the binding information was shuffled between blocks. For scenes in the bound condition, the spatial layouts and binding information were both held constant between blocks. Contextual cueing was absent for unbound scenes with different faces, while contextual cueing was present for unbound scenes with same faces, as well as for both bound scenes with same faces, and bound scenes with different faces. Thus, for scenes with different faces, binding learning was required to facilitate search (Zheng et al., 2019, Experiment 2).

A summary of the discussed studies investigating the learning of binding information can be found in Table 1.

Experiment	Experiment design	Stimulus type	Repeated context	Result	Binding learning
Endo & Takeda (2004), Exp. 1	Dual-phase, Within-subjects	Abstract contours	Configuration repetition	CCE present during transfer epoch	Absent
Endo & Takeda (2004), Exp. 3	Single-phase, Within-subjects		Configuration repetition	CCE present	Absent
Makovski (2016), Experiment 1	Single-phase, Within-subjects	Real-world objects	Location-repeat	CCE absent	Present
Makovski (2017), Experiment 2	Single-phase, Within-subjects		Unbound	CCE absent	Present
Zheng et al. (2019), Experiment 1	Dual-phase, Between-subjects	Faces	Upright-Different faces	CCE absent at start of testing phase	Present
Zheng et al. (2019), Experiment 2	Single-phase, Between-subjects		Unbound	CCE absent	Present
Wan (2019), Experiment 1	Single-phase, Within-subjects	Faces	Repeated-unbound	CCE present	Absent
Wan (2019), Experiment 2	Dual-phase, Within-subjects		Repeated-Different faces	CCE present at start of testing phase	Absent

Table 1. A summary of studies investigating learning of repeated binding information, with abstract contours, real-world objects, and faces.

4. Current Research

We wished to re-examine whether or not identity-to-location binding can be learned. Similar to Zheng et al., (2019), we also used face stimuli to test the learning of the search items' identities, when the items themselves offer little differences in semantic information. Our experiments used within-subjects designs to eliminate possible effects from different populations within the participant pool (for example, in Zheng et al.'s (2019) between-subjects design). Experiment 1 was conducted with a single-phase within-subjects design, with two types of repeated scene conditions: the repeated-unbound condition, where the scenes' respective identity-to-location binding information shuffled between blocks (experimental condition), and the repeated-bound condition, where the scenes' respective identity-to-location binding information remained constant between blocks (control condition). Experiment 2 was conducted with a dual-phase within-subjects design, where select scenes' identity-to-location binding shuffled once—at the end of the first phase. The scene's distractor faces were either all different (experimental condition) or all the same (control condition). The information of Experiments 1 and 2 are included in Table 1.

Our stimulus sets featured upright images of Caucasian faces, based on the predominant ethnicity in Canada, where our studies were conducted. This is consistent with Zheng et al.'s (2019) experiments conducted in China, using Chinese faces. All of the faces in each scene shared the same ethnicity, to minimize search asymmetries from cross-race deficits (Levin, 2000). To avoid directing participants' attention to any one face, and to avoid any extraneous grouping of faces within a scene, we selected faces with the same facial expression and skin tone (Williams, Moss, Bradshaw, & Mattingley, 2005; Glicksohn & Cohen, 2011). We picked faces with neutral expressions as they attract the least attention compared against happy, sad, fearful, angry, and threatening faces (Williams et al., 2005), and do not appear biased toward any one emotion (Horstmann, 2007). Furthermore, the faces in our stimulus set appeared similar to one another, so that if a viewer only compares the most informative facial feature—the eyes (Roberts & Bruce, 1988)—the scene search is still effortful. These considerations were made to increase the search difficulty, and to avoid perceptual pop-out.

EXPERIMENT 1

1. Introduction

We investigated learning in two “repeated” conditions—one with identity-to-location binding, and one without, while scenes in both conditions repeated their respective spatial information. Therefore, each scene in Experiment 1 belonged to one of three conditions: 1) *repeated-unbound* scenes, where the distractor spatial configurations that repeated every block, but the distractor faces shuffled positions each block; 2) *repeated-bound* scenes, where the distractor spatial configurations repeated every block, and the distractor faces’ positions remained constant every block (see Figure 1 for a schematic illustration of the repeated scenes); and 3) *non-repeated* scenes, with distractor spatial configurations that only appeared once throughout the experiment.

We expected reaction times for repeated-unbound scenes to be longer than those of the repeated-bound scenes, due to the constant disruption of identity-to-location binding. If the repeated binding information is learned, then the binding disruption for the repeated-unbound scenes introduces a conflict in learning, which would hinder participants’ search performances. If the repeated binding information is not learned, then the disruptions to the binding information would likely have no effect on participants’ search performances for the repeated-unbound scenes.

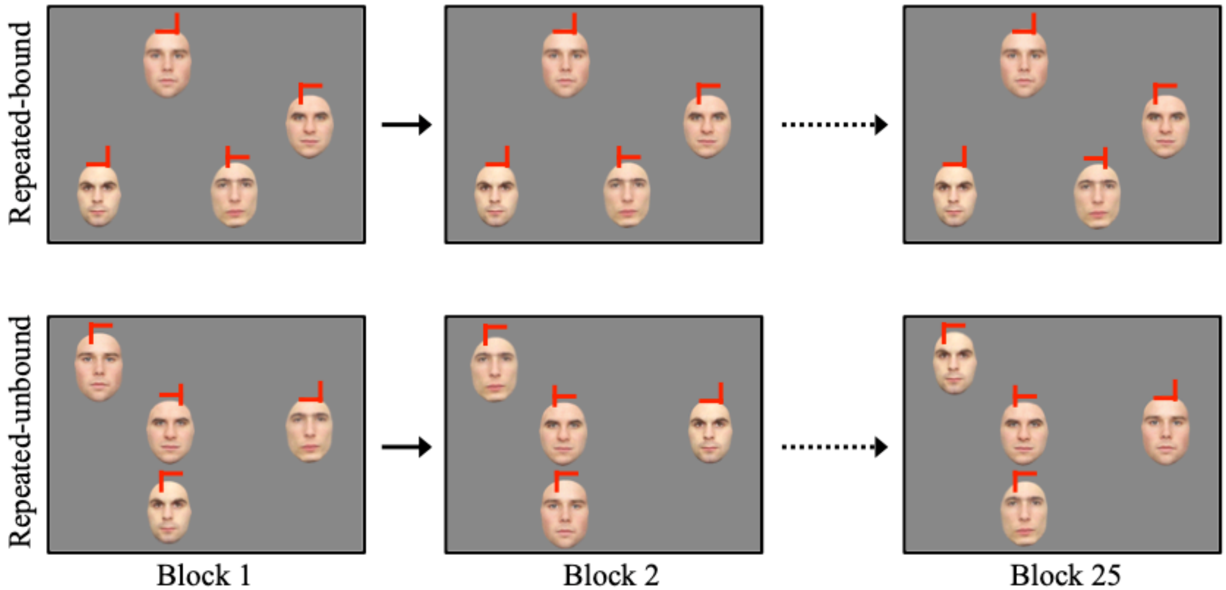

Figure 1. A schematic illustration of the repeated scenes in Experiment 1. The distractor spatial arrangement for repeated-bound scenes (top row) remains constant across blocks. The distractor faces for repeated-unbound scenes (bottom row) shuffles between each block. In both types of repeated scenes, the target remains in the same position. Note: the actual experiment scenes contain 11 distractor faces, for a total of 12 faces in each scene.

2. Method

2.1 Participants

Thirty-six students (M = 18.3, 5 male) from McMaster University's psychology program participated for course credit. All participants had normal or corrected-to-normal vision.

2.2 Materials

The software was programmed in E-Studio 2.0, and run using a 21" Dell monitor (30.76º x 21.53º). Each stimulus consisted of an image of an upright male face (hair and ears removed), beneath a randomly rotated letter 'L' (rotated either 90º or 270º) or a letter 'T' (rotated either 90º or 270º). The faces had an average visual angle of 1.23º x 1.38º, and the letters had an average visual angle of 0.43º x 0.43º. The stimuli in each scene appeared in front of a grey background (138, 138, 138). The fixation cross had a visual angle of 0.67º x 0.67º, and the feedback slide's word span had an average visual angle of 3.10º x 0.57º.

2.3 Procedure

Participants were seated 60 cm away from the screen (unconstrained). On each trial, a centered fixation cross was displayed for 400 ms, followed by a search scene of 11 distractors (a face beneath an 'L') and 1 target (a face beneath a 'T'). The search scene was left on until the participants responded. Upon finding the target, the participants pressed the key corresponding to the pointing direction of the letter 'T's stem. Following the response, a feedback slide was presented for 300 ms. A sample trial is depicted in Figure 2. If the participants answered incorrectly, a bell sound also played from an external speaker. A block consisted of 6 non-repeated trials, 6 repeated-bound trials, and 6 repeated-unbound trials, for a total of 18 trials. In each block, the trials were presented in random sequence. Each non-repeated trial appeared only once for the entire experiment, while repeated-bound and repeated-unbound trials appeared once in each block. Each repeated-bound trial maintained its own spatial configuration (i.e. the same layout of target and distractor positions), where each face appeared in the same position in every block. Each repeated-unbound trial maintained its own spatial configuration, but the distractor faces' positions shuffled every block. The same set of 12 faces was used in each trial of the experiment. In addition to 12 practice trials, participants completed 450 experimental trials (25 blocks, 18 trials per block), for a total of 462 trials. The experiment took 60 minutes to complete.

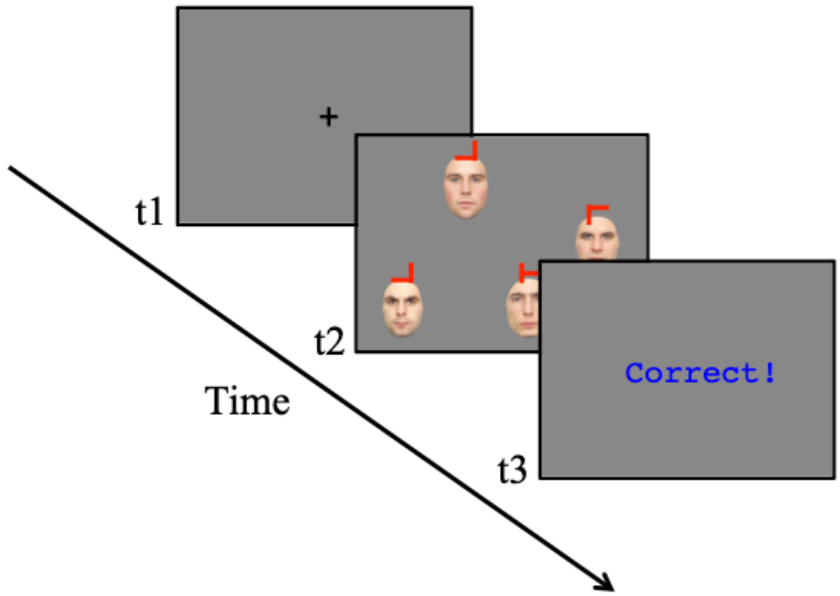

Figure 2. A sample trial in Experiment 1. The fixation cross (t1) appears for 400 ms, followed by the search secne (t2). Following the participant's key press, the feedback slide (t3) appears for 300 ms. Note: the actual experiment scenes contain 11 distractor faces, for a total of 12 faces in each scene.

3. Results

3.1 Data screening

Overall accuracy was above 98.3%. One participant was excluded from analysis due to a high error rate (>10%). Incorrect trials and trials with RTs deviating 2.5 *SD* above and below the mean of each cell were also excluded from analysis, for a total of 6.3% of all trials. Figure 3 illustrates the mean RTs as a function of epoch (defined here as a group of five consecutive blocks) and of scene condition. All *p*-values were Bonferroni-corrected to maintain a familywise α of 0.05.

3.2 RT analysis

A repeated-measures ANOVA with factors configuration (non-repeated, repeated-bound, and repeated-unbound) and epoch (1–5, each consisting of five consecutive blocks) was conducted. The analysis revealed a significant main effect of configuration, $F(2, 68) = 29.19$, $p < 0.001$, $\eta_p^2 = 0.52$, and epoch, $F(4, 136) = 16.26$, $\hat{\varepsilon} = 0.66$, $p < 0.001$, $\eta_p^2 = 0.55$. The interaction between configuration and epoch was not significant, $F(8, 272) = 1.27$, $\hat{\varepsilon} = 0.80$, $p = 0.26$, $\eta_p^2 = 0.29$. A post-hoc Tukey Honest Significant Difference (HSD) test revealed that the repeated-bound RTs (*M* = 2231.41 ms, *SD* = 936.58 ms) were significantly longer than both the repeated-unbound RTs (*M* = 1882.18 ms, *SD* = 898.17 ms) and the non-repeated RTs (*M* = 2044.19 ms, *SD* = 920.68 ms), *p*s < 0.05. The difference in RTs between repeated-unbound scenes and non-repeated scenes was not significant, $p = 0.07$.

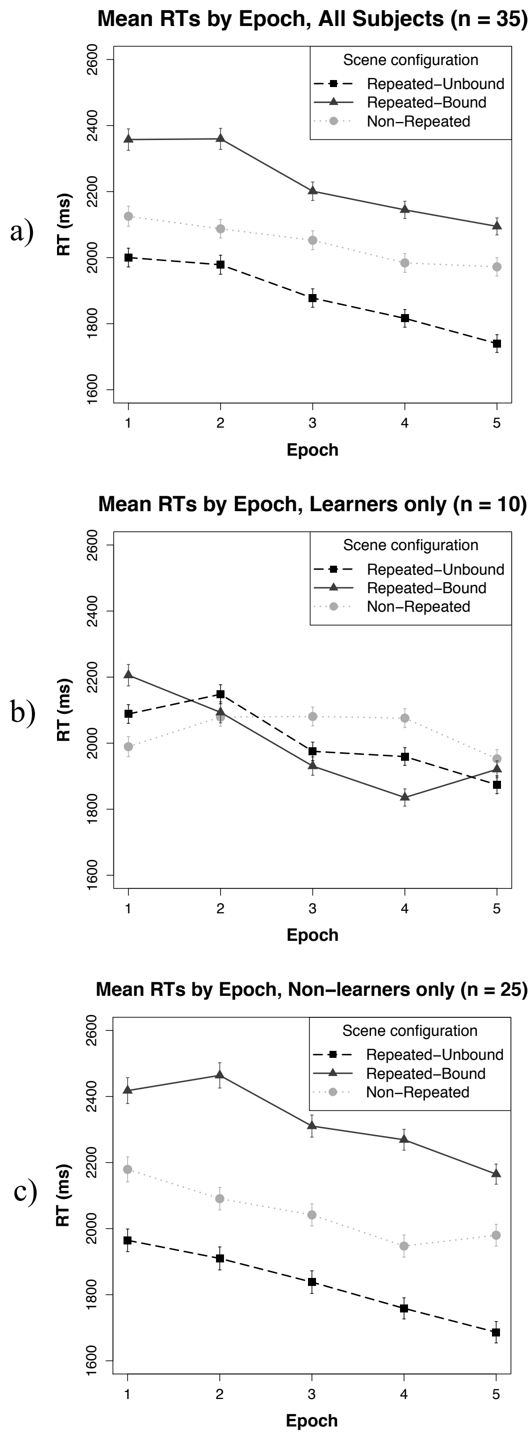

Figure 3. Mean RTs for Experiment 1 grouped by epoch and scene condition for a) all participants, b) learners only, and c) non-learners only. Participants with a positive CCE for repeated-bound scenes, averaged over Epochs 3–5, were designated as “learners.” Participants with zero or a negative CCE for repeated-bound scenes, averaged over Epochs 3–5, were designated as “non-learners.” Error bars represent ±1 *SE* from the mean.

As discussed in the General Introduction, not all participants may have successfully learned the repeated contexts in this paradigm, due to possible individual differences in learning performance. The participants were separated into two groups: "learners"—participants with a positive CCE for repeated-bound scenes, averaged over Epochs 3–5 ($n = 10$), and "non-learners"—participants with a negative CCE for repeated-bound scenes, averaged over Epochs 3–5 ($n = 25$).

Analyses were conducted separately for learners and non-learners, with the same 3 x 5 repeated-measures design. For learners, the analysis revealed a significant main effect of epoch, $F(4, 36) = 2.96$, $p < 0.05$, $\eta_p^2 = 0.58$. The main effect of scene condition was not significant, $F(2, 18) = 0.13$, $p = 0.88$, $\eta_p^2 = 0.08$. The interaction between scene condition and epoch was significant, $F(8, 72) = 2.10$, $p < 0.05$, $\eta_p^2 = 0.91$. The simple main effect of scene condition was not significant at all epochs, all ps > 0.25. A Tukey HSD test revealed no significant pairwise difference in RTs for the repeated-bound scenes ($M = 1996.02$ ms, $SD = 881.47$ ms), the repeated-unbound scenes ($M = 2008.85$ ms, $SD = 872.51$ ms), and the non-repeated scenes ($M = 2035.55$ ms, $SD = 904.36$ ms), all ps > 0.95.

For non-learners, the analysis revealed significant main effects of epoch and scene condition, $F(4, 96) = 13.41$, $p < 0.001$, $\hat{\varepsilon} = 0.63$, $\eta_p^2 = 0.65$; $F(2, 48) = 75.61$, $p < 0.001$, $\eta_p^2 = 0.84$, respectively. The interaction between scene condition and epoch was not significant, $F(8, 192) = 1.52$, $p = 0.15$, $\eta_p^2 = 0.53$.

3.3 CCE analysis over Epochs 3–5

The magnitude of the CCEs for repeated-bound scenes (mean RT[non-repeated scenes] - mean RT[repeated-bound scenes] and CCEs for repeated-unbound scenes (mean RT[non-repeated scenes] - mean RT[repeated-unbound scenes]) were averaged over Epochs 3 to 5 and were compared for all subjects, learners, and non-learners. Figure 4 illustrates the mean CCEs for both repeated scene types, separated by subject group. The analysis revealed an overall significant difference in mean CCE between repeated scene types, $t(64) = -5.99$, $p < 0.001$. Although bound repeated scenes led to a slightly greater CCE compared to unbound scenes, the difference in mean CCE between the two repeated scene types was not significant for learners, $t(18) = -0.38$, $p > 0.5$. Interestingly, for non-learners, the mean CCE for repeated-unbound scenes was significantly greater than the mean CCE for repeated-bound scenes, $t(48) = -9.21$, $p < 0.001$.

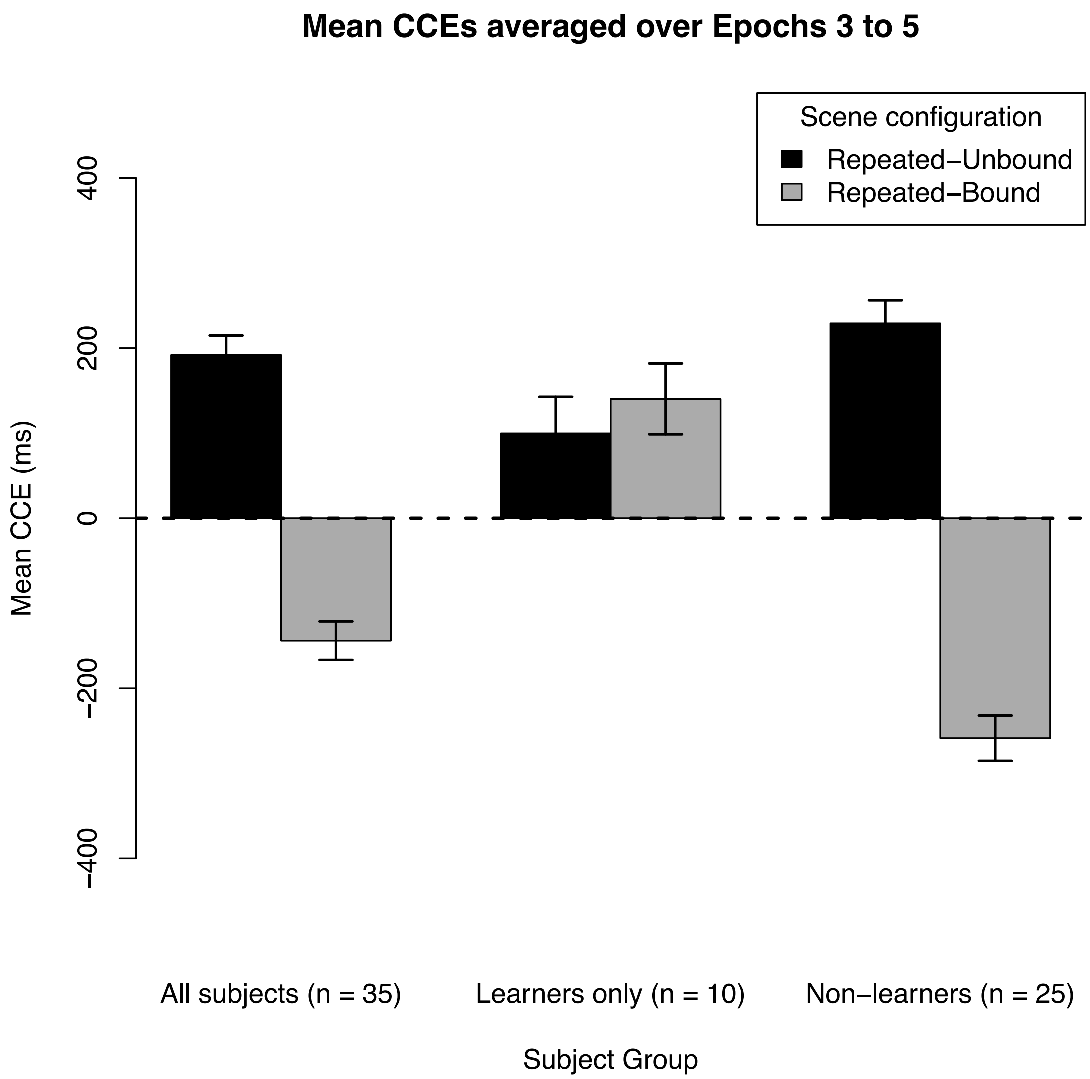

Figure 4. Mean CCEs for repeated-bound and repeated-unbound scenes averaged over Epochs 3 to 5 for Experiment 1, organized by subject group. Participants with a positive CCE for repeated-bound scenes, averaged over Epochs 3–5, were designated as "learners." Participants with zero or a negative CCE for repeated-bound scenes, averaged over Epochs 3–5, were designated as "non-learners." Error bars represent ±1 *SE* from the mean.

4. Discussion

4.1 Negative contextual cueing for repeated-bound scenes

Reaction times for Experiment 1 improved across epochs for all scene conditions, with a persistent CCE for repeated-unbound scenes. We expected to observe the largest CCE for repeated-bound scenes, since the repeated-bound scenes kept the most information constant (the spatial layout and the identity-to-location binding). Moreover, we expected to observe a smaller or no CCE for repeated-unbound scenes, since the repeated-unbound scenes shuffled their identity-to-location binding every block. However, the overall reaction times for repeated-bound scenes were consistently longer than the reaction times for non-repeated scenes. Only ten participants exhibited a positive CCE for repeated-bound scenes, satisfying our definition of being a "learner." Nevertheless, the main effect of scene condition on reaction time in the learners-only analysis approached significance ($p = 0.08$). While this experiment may have had insufficient learners for the effect to reach significance, the repeated contexts might not have facilitated search for learners as strongly as in other studies. For learners, the search facilitation for repeated-bound scenes may have been made stronger if the learners completed more experimental blocks (considering the significant interaction between scene condition and epoch), potentially increasing the search advantage for repeated-bound scenes over non-repeated scenes. Overall search facilitation may have been weaker due to the similarities between repeated-bound and the repeated-unbound contexts.

4.2 Positive contextual cueing for repeated-unbound scenes

Experiment 1 featured a restrictive stimulus set as in Zheng et al. (2019), rendering the two repeated contexts highly similar. Both the repeated-bound and repeated-unbound scenes used the same set of faces and letters, but only the repeated-bound scenes kept their respective identity-to-location binding constant. It is possible that for the repeated-unbound scenes, the non-learners were able to ignore the shuffling faces, while still being able to take advantage of the repeated configurations. In contrast, for the repeated-bound scenes, it is possible that the non-learners unconsciously processed the binding of local identity-to-location in the repeated-bound scenes, inadvertently introducing an extra learning demand, consequently leading to a cost in CC.

Moreover, in a within-subjects design, strategy or processes from one condition may influence those from another condition. For example, the presence of the repeated-unbound condition might reduce the learning of identity information in the repeated-bound condition. Viewing both types of repeated scenes could have introduced a conflict in search facilitation, as only the repeated-bound scenes' identity-to-location binding remained unchanged in every block. This conflict may have driven a global slowdown in search, as observed in Vaskevich & Luria (2018), where search times for non-repeated scenes alone were faster than search times for non-repeated scenes when presented with repeated scenes. For Experiment 1's non-learners, the conflict in search facilitation affected CCEs for repeated-bound scenes. In contrast, Experiment 1's learners may have learned the repeated-bound scenes and repeated-unbound scenes such that the two repeated contexts did not conflict with one another. The lack of conflict could have been driven by the learners weighing spatial information over identity-to-location binding, or by the learners ignoring the identity-to-location binding altogether. In both cases, learners would have perceptually treated both types repeated scenes as the same condition. This is supported by the non-significant difference in RTs between repeated-bound and repeated-unbound scenes. Therefore, the learners in Experiment 1 may have implicitly re-weighted the learned repeated contexts by judging certain information as irrelevant or redundant, and subsequently ignoring it.

EXPERIMENT 2

1. Introduction

In Experiment 1, we observed that the inconsistent pairing of repeated identities to repeated locations hindered learning. In Experiment 2, we investigated the possibility of viewers having learned binding information when given the opportunity to learn. We implemented two phases: a learning phase and a testing phase. During the learning phase, the repeated scenes' respective identity-to-location bindings were kept intact. During the testing phase, the repeated scenes maintained their respective spatial layouts from the learning phase, but each repeated scene adopted a new identity-to-location binding (see Figure 5 for a schematic illustration of both types of repeated scenes). Thus, each repeated scene's binding of identity-to-location was shuffled once—between the end of the learning phase and the start of the testing phase. We tested—following the expected learning of identity-to-location binding information—whether or not a change to binding would lead to a cost in the contextual cueing effect.

Similar to Experiment 1, the stimulus set consisted of 12 Caucasian faces: 11 distractor male faces, and one target female face. Instead of searching for a target *T* among distractor *L*'s as in Experiment 1, the task was to search among the faces themselves. In doing so, the faces became task-relevant. All of the selected faces had similar relative distances between their respective facial features, making it difficult to make quick or rough judgments on the sex of each face (Burton, Bruce, & Dench, 1993).

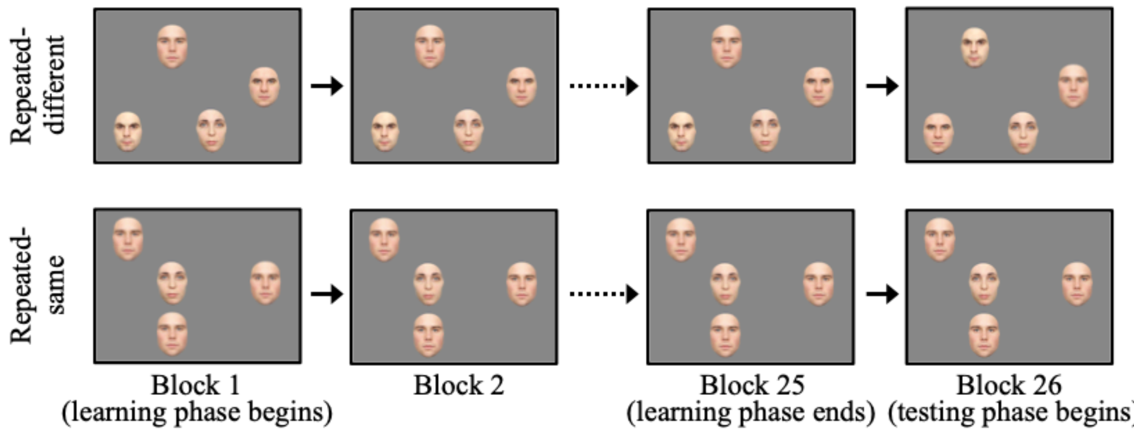

Figure 5. A schematic illustration of the repeated scenes in Experiment 2. During the learning phase (Block 1 to Block 25), the distractor spatial arrangement and identity-to-location binding for repeated-different scenes (top row) remains constant. During the testing phase (Block 26 to Block 30), the distractor faces in the repeated-different scenes shuffle identity-to-location binding but keep the same spatial arrangement. The spatial arrangement and identity-to-location binding for repeated-same scenes (bottom row) is held constant throughout the learning phase and the testing phase. In both types of repeated scenes, the target remains in the same position. Note: the actual experiment scenes contain 11 distractor faces, for a total of 12 faces in each scene.

One half of the search arrays consisted of non-repeated scenes that appeared only once each. The other half of the search arrays consisted of repeated scenes that preserved the spatial layout of target and distractors with fixed identity-to-location binding every block (with the exception of a one-time binding shuffle between the learning phase and the testing phase). For both non-repeated and repeated conditions, half of the scenes' distractors were identical copies of one face (*same faces* condition), and the other half of the scenes' distractors were all different (*different faces* condition). Search performance for *repeated-different faces* scenes was contrasted against that of *non-repeated-different faces* scenes, and that of *repeated-same faces* scenes, from which we expected the weakest and strongest search performances, respectively.

While we expected that CC would manifest for both repeated-different faces scenes and repeated-same faces scenes (such that binding is learned in the learning phase), we expected that the one-time binding shuffle would only eliminate the contextual cueing for the repeated-different faces scenes. The shuffling of identity-to-location binding would affect the appearances of repeated scenes with different faces, but not those of repeated scenes with same faces.

2. Method

2.1. Participants

Fifty-four students (M = 18.8, 14 male) from McMaster University's psychology program participated for course credit. All participants had normal or corrected-to-normal vision.

2.2. Stimuli

The software was programmed in PsychoPy (version 1.90.2), and run using a 21" Dell monitor (30.76º x 21.53º). Each stimulus consisted of an image of an upright face (hair and ears removed). The faces had an average visual angle of 1.23º x 1.38º. The stimuli in each scene appeared in front of a grey background (138, 138, 138). The fixation cross had a visual angle of 0.95º x 0.95º, and the feedback slide's word span had an average visual angle of 7.01º x 1.15º.

2.3. Procedure

Participants were seated 60 cm away from the screen (unconstrained). On each trial, a centered fixation cross was displayed for 200 ms, followed by a search scene of 11 distractor male faces, and 1 target female face. The search scene was left on until the participants responded. Upon finding the target, the participants pressed the spacebar, at which point the faces in the scene were replaced with numbers from 1 to 12. The participants then pressed the key on the keyboard labeled with the number corresponding to the target's position. Following the response, a feedback slide was presented for 300 ms. A sample trial is depicted in Figure 6. A block consisted of 24 trials, with 12 scenes that appeared in every block (repeated, "old" scenes), and 12 scenes that only appeared once during the experiment (non-repeated, "new" scenes). Each repeated scene maintained the spatial arrangement of its target and distractor positions, while keeping binding of identity-to-location constant.

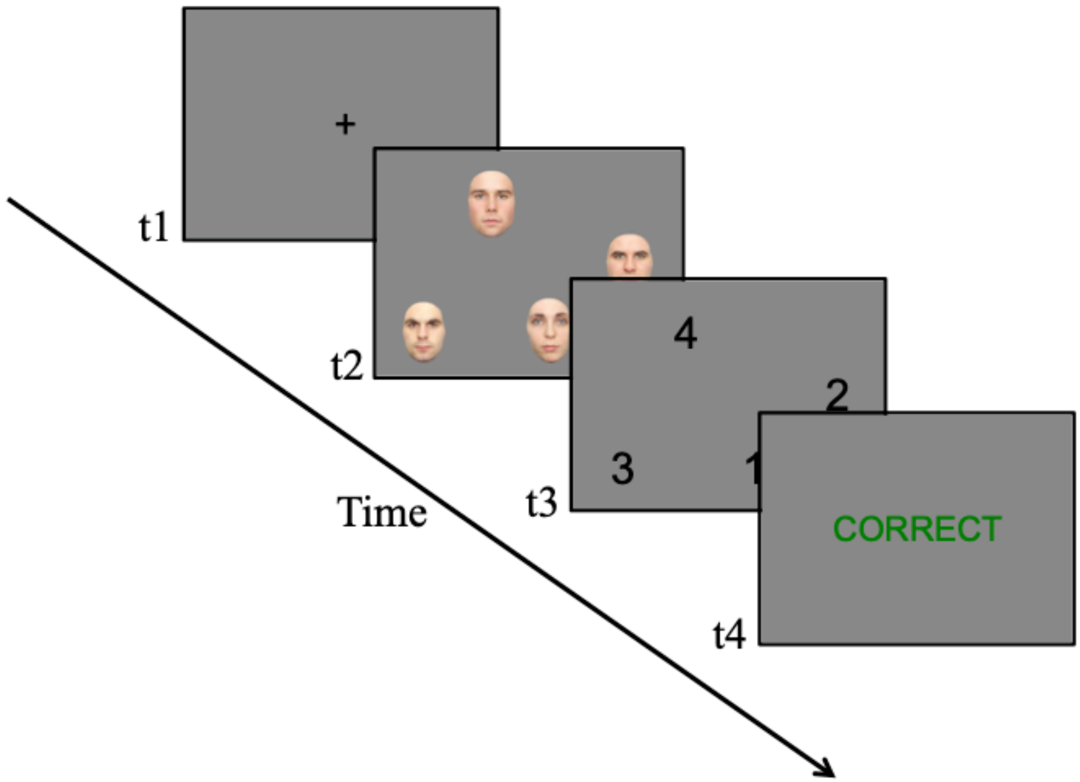

Figure 6. A sample trial in Experiment 2. The fixation cross (t1) appears for 200 ms, followed by the search scene (t2). The number slide (t3) appears following a spacebar press. Following the subject's response, a feedback slide (t4) appears for 300 ms. Note: the actual experiment scenes contain 11 distractor faces, for a total of 12 faces in each scene.

The distractor faces in half of the non-repeated and repeated scenes were identical copies of one male face ("same faces" condition), and the distractor faces in the remaining non-repeated and repeated scenes were different male faces ("different faces" condition). The same distractor face appeared in each scene within the same faces condition, and the same 11 distractor faces appeared in each scene within the different faces condition.
Participants completed 10 practice trials, followed by a learning phase (25 blocks) and a testing phase (5 blocks). For the testing phase, the distractor faces' positions for all of the repeated scenes were shuffled, but they maintained the same spatial arrangement. This identity-to-location binding was shuffled once, while maintaining their respective distractor and target positions. After the testing phase, the participants were asked whether or not they noticed any patterns or any repeated elements. The experiment took 60 minutes to complete.

3. Results

3.1. Data screening

Overall accuracy was above 97.5%. Four participants were excluded from analysis (one due to data corruption, and three due to a high error rate (>10%)). Incorrect trials and trials with RTs deviating 2.5 *SD* above and below the mean of each cell were also excluded from analysis, for a total of 8.3% of all trials. Figure 7 displays the mean RTs as a function of epoch (defined here as a group of five consecutive blocks) and configuration–face condition pair, for all subjects, learners, and non-learners. All *p*-values were Bonferroni-corrected to maintain a familywise α of 0.05.

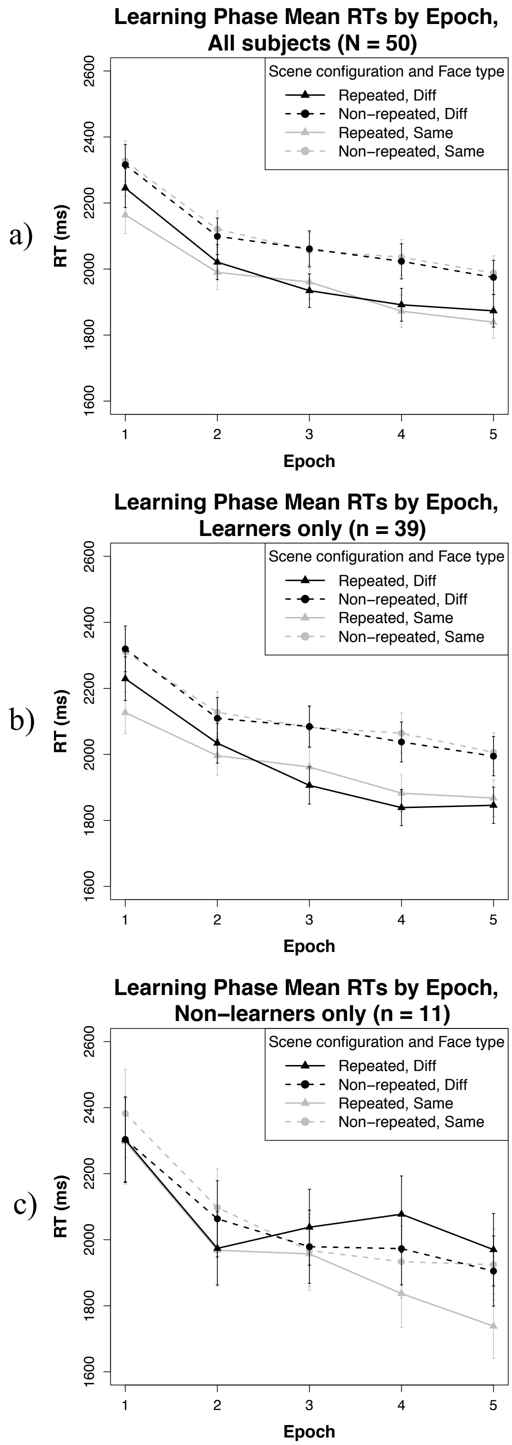

Figure 7. Learning phase mean RTs for Experiment 2 grouped by epoch and scene condition for a) all subjects, b) learners only, and c) non-learners only. Participants with a positive CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "learners." Participants with zero or a negative CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "non-learners." Error bars represent ±1 *SE* from the mean.

3.2. RT analysis

Mean RTs for the learning phase were submitted to a repeated-measures ANOVA with factors configuration (non-repeated vs. repeated), face (different vs. same), and epoch (1–5, each consisting of five consecutive blocks). The analysis revealed significant main effects of configuration, $F(1, 49) = 92.72$, $p < 0.01$, $\eta_p^2 = 0.65$, and epoch, $F(4, 196) = 80.35$, $p < 0.01$, $\hat{\varepsilon} = 0.65$, $\eta_p^2 = 0.87$. The main effect of face was not significant, $F(1, 49) = 0.48$, $p = 0.68$, $\eta_p^2 = 0.0003$. The two-way interactions were not significant (all p's > 0.14). The three-way interaction was not significant, $F(4, 196) = 0.69$, $p > 0.5$, $\hat{\varepsilon} = 1.01$, $\eta_p^2 = 0.08$.

As in Experiment 1, not all participants may have successfully learned the repeated contexts in this paradigm, due to possible individual differences in learning performance. The participants were separated into two groups: "learners"—participants with a positive CCE for different-faces scenes, averaged over Epochs 3–5 ($n = 39$), and "non-learners"—participants with a negative CCE for different-faces scenes, averaged over Epochs 3–5 ($n = 11$).

Analyses were conducted separately for learners and non-learners, with the same 2 x 2 x 5 repeated-measures design. For learners, the analysis revealed significant main effects of configuration and epoch, $F(1, 38) = 125.32$, $p < 0.01$, $\eta_p^2 = 0.76$; $F(4, 152) = 23.42$, $p < 0.01$, $\hat{\varepsilon} = 0.67$, $\eta_p^2 = 0.68$, respectively. The main effect of face was not significant, $F(1, 38) = 0.01$, $p > 0.5$, $\eta_p^2 = 0.0001$. The two-way interactions were not significant (all p's > 0.25). The three-way interaction was not significant, $F(4, 152) = 1.06$, $p > 0.5$, $\eta_p^2 = 0.20$.

For non-learners, the analysis revealed a significant main effects of scene and epoch, $F(4, 40) = 23.65$, $p < 0.001$, $\hat{\varepsilon} = 0.65$, $\eta_p^2 = 0.86$, respectively. The main effects of configuration and face were not significant, $F(1, 10) = 3.75$, $p = 0.24$, $\eta_p^2 = 0.28$; $F(1, 10) = 1.05$, $p > 0.5$, $\eta_p^2 = 0.10$, respectively. The configuration x face interaction approached significance, $F(1, 10) = 7.44$, $p = 0.06$, $\eta_p^2 = 0.43$. The other two-way interactions were not significant (p's > 0.13). The three-way interaction was not significant, $F(4, 40) = 0.87$, $p > 0.5$, $\eta_p^2 = 0.31$.

3.3. Magnitude of CCE over Learning Phase Epochs 3–5

A separate analysis was conducted to compare the CCEs for same-faces scenes and the CCEs for different-faces scenes. The mean CCEs averaged over Epochs 3–5 for each subject group (all subjects, learners only, and non-learners only) are illustrated in Figure 8. The overall difference between CCEs for same-faces scenes and CCEs for different-faces scenes was not significant, $t(98) = 0.19$, $p > 0.5$. For learners, the difference between CCEs for same-faces scenes and CCEs for different-faces scenes was not significant, t(76) = -1.57, $p = 0.36$. For non-learners, the difference between CCEs for same-faces scenes and CCEs for different-faces scenes was significant, $t(20) = 4.04$, $p < 0.01$.

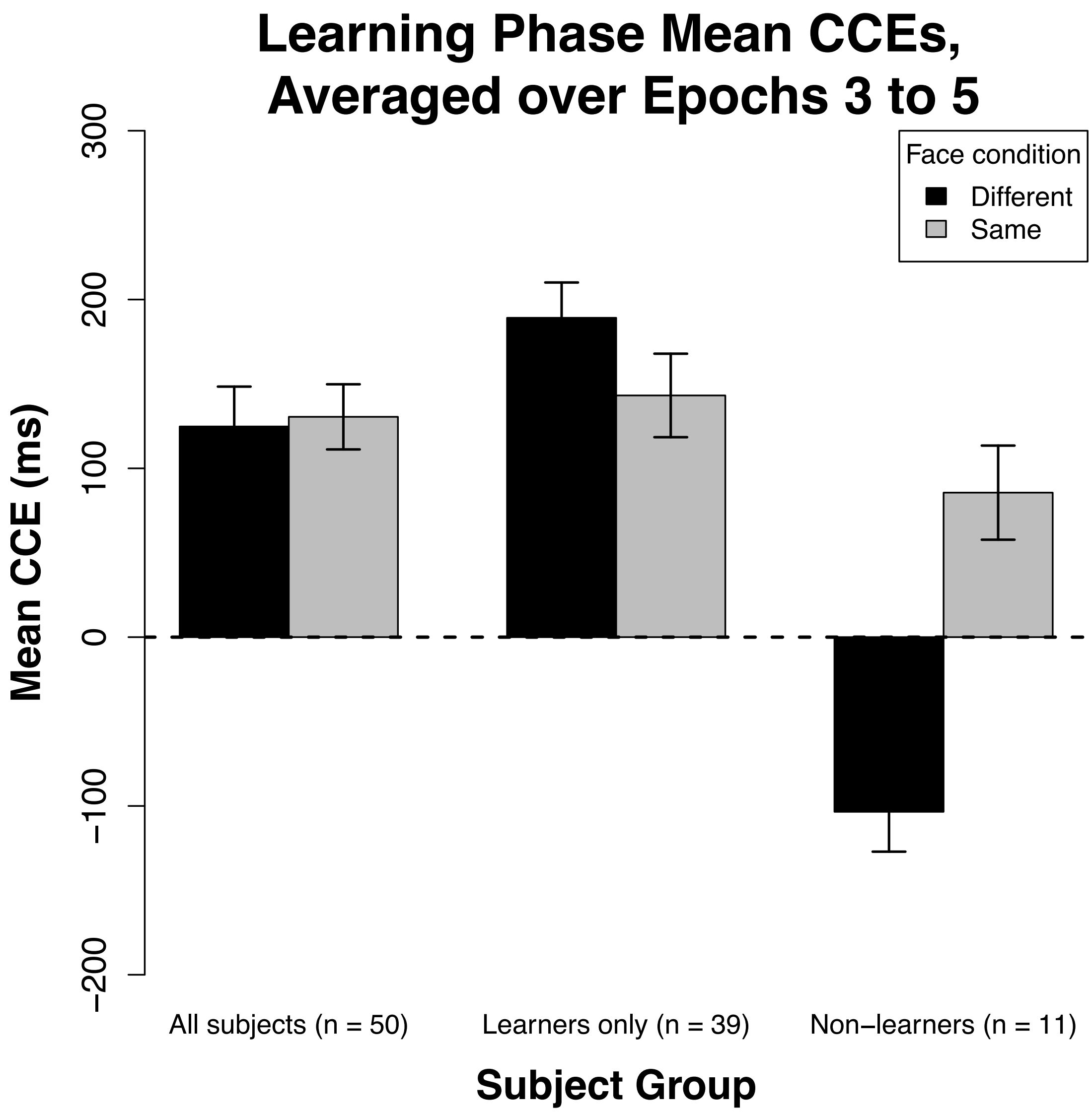

Figure 8. Mean CCEs for same-faces scenes and different-faces scenes averaged over Epochs 3 to 5 for Experiment 2, organized by subject group. Participants with a positive CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "learners." Participants with zero or a negative CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "non-learners." Error bars represent ±1 *SE* from the mean.

3.4. End of learning phase vs. start of testing phase analysis

A planned repeated-measures ANOVA with factors face (same vs. different) and phase (learning phase vs. testing phase: Blocks 24–25 vs. Blocks 26–27) was conducted to investigate the change in CCEs following the one-time shuffle of identity-to-location binding. We only selected data from two blocks because of the possible learning in the subsequent blocks for the new set of repeated contexts during the testing phase. The analysis was completed for each subject group (all subjects, learners only, and non-learners only). Mean CCEs for the end of the learning phase (Blocks 24–25) and the start of the testing phase (Blocks 26–27) are illustrated in Figure 9, organized by face condition and subject group.

For all subjects, the main effects of face and phase were not significant, $F(1, 49) = 0.01$, $p > 0.5$; $F(1, 49) = 2.39$, $p = 0.36$, respectively. The interaction between face and phase was not significant, $F(1, 49) = 2.60$, $p = 0.11$. For learners, the main effect of phase approached significance, $F(1, 38) = 5.54$, $p = 0.07$. The main effect of face was not significant, $F(1, 38) = 0.20$, $p > 0.5$. The interaction between face and phase was not significant, $F(1, 38) = 0.27$, $p > 0.5$. For non-learners, the main effects of face and phase were not significant, $F(1, 10) = 1.22$, $p = 0.29$; $F(1, 10) = 0.36$, $p > 0.5$, respectively. The interaction between face and phase was marginally significant, $F(1, 10) = 8.47$, $p = 0.046$.

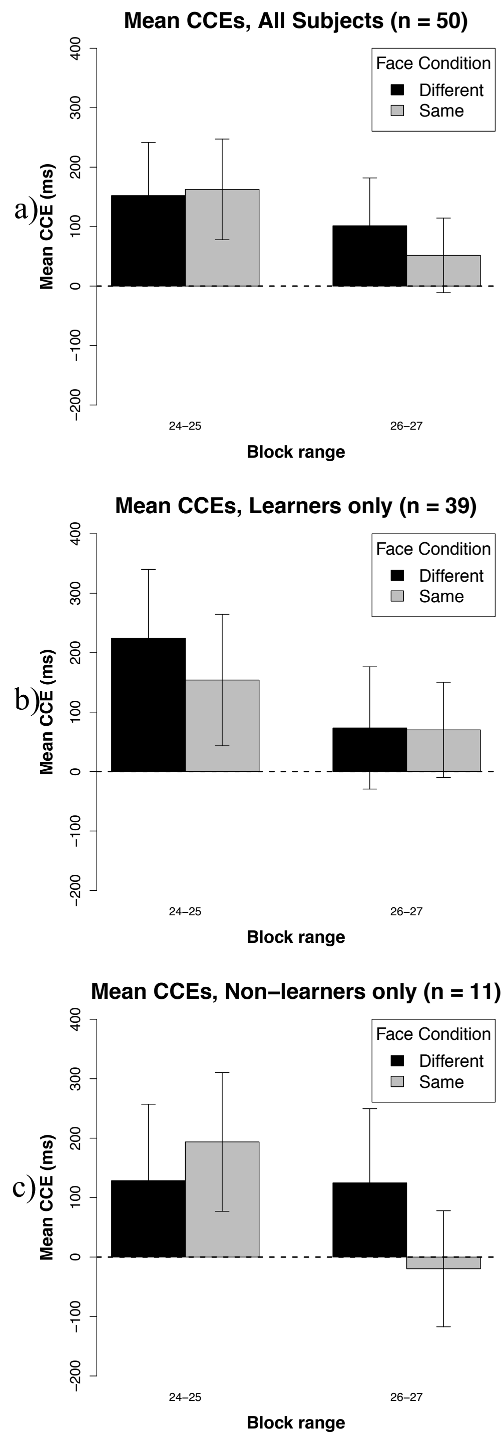

Figure 9. Mean CCEs for Experiment 2 at the end of learning phase (Blocks 24–25) and at the start of testing phase (Blocks 26–27), for a) all subjects, b) learners only, and c) non-learners only. Participants with a positive CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "learners." Participants with zero or a negative CCE for repeated-different scenes, averaged over Epochs 3–5, were designated as "non-learners." Error bars represent ±1 *SE* from the mean.

4. Discussion

4.1 Weak identity learning during the learning phase

During the learning phase, the learners' reaction times for repeated scenes were consistently shorter than reaction times for non-repeated scenes, for both same or different face conditions. At the beginning of the testing phase, the contextual cueing effects for both face groups incurred a cost. This cost could have been due to a time gap during the experiment between the learning phase and the testing phase. Nevertheless, the difference between CCEs for same-faces scenes and CCEs for different-faces scenes is more important. When comparing the end of the learning phase and the beginning of the testing phase, the cost to the contextual cueing effect was slightly larger for different faces compared to same faces, but the difference was not significant. This suggests that subjects learned little about the identity information from the different-faces scenes during the learning phase.

4.2 Stimulus set complexity and identity-to-location learning

The lack of statistically significant differences in CCE between the different-faces and same-faces scenes conflicts with previous studies. Experiment 2's search arrays consisted of faces, where sex was the only semantic category by which the target and distractors can be distinguished. In contrast, Makovski's (2016; 2017) tasks required subjects to make more complicated categorical comparisons in order to locate their targets, as their search arrays consisted of images of real-world objects, belonging to several different semantic categories. If the discrepancy in stimulus set complexity made the search task in Experiment 2 less effortful than search tasks with real-world objects, then the nature of the participants' learning in Experiment 2 may also have differed from that of Makovski's (2016; 2017) experiments. While participants in Makovski's (2016; 2017) experiments may have learned identity-to-location binding information due to viewing complex stimuli, the contextual cueing effects observed in Experiment 2 suggest that the participants primarily learned the repeated spatial configurations. The repeated identity-to-location binding may have either been learned to a lesser degree, or ignored. For the participants' learning, the lesser weighting of identity-to-location binding information is reflected in the non-significant difference in CCE change between scenes with different faces and scenes with same faces.

4.3 The nature of the learned information

4.3.1 Learning phase duration and the nature of learned information

Experiment 2's learning phase was five blocks longer than Zheng et al.'s (2019) learning phase. We expected that the longer learning phase would have contributed to a stronger learning of repeated identity-to-location binding information. For learners, CCE for both same-faces and different-faces scenes was positive throughout the learning phase. We expected that the learners' CCEs established during the learning phase would be eliminated, following the shuffling of identity-to-location binding information at the start of the testing phase. Instead, the CCEs remained positive throughout the testing phase, with no significant decrease in magnitude.

To eliminate a participant's CCE, the repeated information that promoted its establishment must be disrupted. But the CCE will not be eliminated unless the participant's learning sufficiently depends on the repeated information. In Zheng et al. (2019), the one-time shuffle of identity-to-location binding in the testing phase disrupted the contextual cueing effect, but CC quickly re-emerged and persisted until the end of the experiment. This suggested that participants' learning in Zheng et al. (2019) included the repeated identity-to-location binding as a major component. For participants in Experiment 2, the repeated identity-to-location binding did not appear to be a major component in their learning.

The longer learning phase in Experiment 2 may have provided the participants sufficient extra practice to adjust their search strategies, favouring spatial information over identity-to-location binding. The inclusion of the same faces condition may have also led participants to implicitly favour learning repeated spatial information, as only spatial information varied between scenes in this condition. If the participants initially learned the repeated scenes' spatial configurations and identity-to-location binding with equal weighting, then part of their later search facilitation may have involved implicitly deciding that identity-to-location binding was redundant.

4.3.2 Presented repeated contexts and the nature of learned information

The idea of re-weighting learned repeated information assumes that participants can successfully learn all types of presented repeated information. However, if one establishes contextual cueing for repeated scenes with multiple repeated contexts, not all of the repeated contexts may have necessarily been learned. For example, although participants established CC for scenes with abstract contours when both configuration and identity information were repeated, CC for the same scenes persisted when only the location information was preserved, while CC for the same scenes was eliminated when only the identity information was preserved (Endo & Takeda, 2004). This suggests that for scenes with abstract contours, participants only learned the repeated location information despite both location and identity information being repeated together. For real-world objects, participants successfully established CC for the "all-repeat" scenes, which preserved identity-to-location binding (Makovski, 2017). Although both the "all-repeat" and "unbound" scenes repeated location and identity information, the randomized identity-to-location binding in the "unbound" scenes may have interfered with participants' learning, preventing the establishment of CC.

It is possible that the learners in Experiment 2 simply did not learn identity-to-location information from the repeated different-faces scenes. In Experiment 2's learning phase, the repeated scenes held location information and identity-to-location binding information constant. Most participants successfully established contextual cueing during the learning phase, and their CCEs persisted during the beginning of the testing phase (where only the repeated location information remained constant). Therefore, the participants' CCEs may have been driven exclusively by the repeated location information. Despite both repeated contexts were presented together during the learning phase, the participants may have only learned the repeated location information.

GENERAL DISCUSSION

1. Identity-to-location binding as a repeated context

1.1 Overall contextual cueing effects for repeated scenes

We used scenes of faces to explore how changes to identity-to-location binding can affect contextual cueing. The manipulation to identity-to-location binding in Experiment 1 occurred throughout the experiment, where identity-to-location binding for repeated-unbound scenes shuffled every block, while the identity-to-location binding for repeated-bound scenes remained constant across blocks.

In Experiment 1, we expected to observe positive CC for repeated-bound scenes, and weaker CC for repeated-unbound scenes. Instead, we observed an overall negative CCE for repeated-bound scenes, and an overall positive CCE for repeated-unbound scenes. Each block in Experiment 1 included repeated-bound scenes, repeated-unbound scenes, and non-repeated scenes, all of which shared the same stimulus set. The appearance of this restrictive stimulus set in each scene condition may have interfered with learning the fixed identity-to-location binding information from the repeated-bound scenes. We contrast this finding with Makovski (2017), where they observed that individuals display positive contextual cueing for scenes that fixed identity-to-location binding, and no contextual cueing for scenes that shuffled identity-to-location binding. Their experiments featured a large stimulus set of images of real-world objects, belonging to multiple semantic categories (Makovski, 2017), while our experiments featured a restrictive stimulus set of faces, where each face shared the same ethnicity, facial expression, and similar facial features.

In Experiment 2, the identity-to-location binding shuffled once for repeated scenes, between the learning phase and the testing phase. Throughout the learning phase, the identity-to-location binding for each repeated scene was held constant. At the testing phase, the identity-to-location binding for each repeated scene was shuffled, and then held constant for the remainder of the experiment.

In Experiment 2, we expected to observe positive CC for repeated-different faces scenes during the learning phase, followed a sharp decrease in or loss of CC at the start of the testing phase, and the re-emergence and persistence of CC for the remainder of the testing phase. We also expected CCEs for repeated-same faces scenes to remain positive throughout the experiment, as the one-time shuffling of identity-to-location binding only affected repeated-different faces scenes. We observed an overall positive CCE in Experiment 2's learning phase and testing phase, both for repeated-different faces scenes and for repeated-same faces scenes. At the start of the testing phase, CCEs for both repeated-different faces scenes and repeated-same faces scenes decreased slightly, but did not disappear completely. This may suggest that the repeated identity-to-location binding information in the different faces condition did not contribute to the subjects' learning. Alternatively, some subjects may have learned the repeated identity-to-location binding information during the learning phase, but eventually weighted repeated spatial information over repeated identity-to-location binding information. Subjects may have gradually perceived the faces in the different faces condition to appear similar to each other. Taken together, if the stimuli within a repeated scene appear sufficiently similar to one another (e.g. in colour, in shape, or in semantic category, etc.), then repeating identity-to-location binding information may not be necessary for contextual cueing.

1.2 Contextual cueing effects for learners

We performed additional analyses for each experiment by separating each experiment's participants into two populations (learners and non-learners), based on their performances in their respective search task. Ten participants in Experiment 1 established a mean positive contextual cueing effect for repeated-bound scenes (learners) between Epochs 3–5. These participants also successfully established a mean positive contextual cueing effect for repeated-unbound scenes between Epochs 3–5. The learners in Experiment 1 may have employed a search strategy that prioritized learning spatial information, or they may simply have not experienced any interference in learning the fixed identity-to-location binding information. The difference between the learners' mean CCE for repeated-bound scenes and the learners' mean CCE for repeated-unbound scenes was not significant. This suggests that repeated identity-to-location binding information may not contribute to learning when scenes that preserve binding and scenes that shuffle binding appear together.

In Experiment 2, 32 participants established a mean positive contextual cueing effect for different-faces scenes (learners) between Epochs 3–5. These participants also successfully established a mean positive contextual cueing effect for same-faces scenes. The difference between the learners' mean CCE for different-faces scenes and the learners' mean CCE for same-faces scenes was not significant, which suggests that viewers can successfully learn repeated spatial information when their search involves a sex selection task.

2. Learner vs. non-learner designation

We observed a low proportion of learners in Experiment 1 (10 learners out of 35 participants), but observed a larger proportion of learners in Experiment 2 (32 learners out of 50 participants). Studies in the literature have also recorded varying frequencies of positive and negative CCE over the last three epochs of their respective experiments. In Lleras & von Mühlenen (2004)'s third experiment, 7 out of 20 participants had a mean positive CCE when instructed to actively search for their target. When instructed to let their target pop out, 16 out of 20 participants had a mean positive CCE (Lleras & von Mühlenen, 2004). Kunar & Wolfe (2011) observed 11 out of 16 participants achieving a positive CCE for repeated scenes with multiple possible target locations. Due to individual differences in learning, some viewers will successfully learn the repeated information within a given visual search paradigm, but others will not. Therefore, the nature of the search task and the experiment design can affect the proportion of learners and non-learners, in addition to the set learner vs. non-learner criterion.

Our learner vs. non-learner criterion was based off of Chun & Jiang (1998), where they averaged individual CCEs over the final three epochs of their experiment. We also designated participants as learners or non-learners based on the sign of their respective CCE, similar Lleras and von Mühlenen (2004)'s characterizations of their subjects' contextual cueing (positive and negative). The differences in mean RTs between learners and non-learners in Experiments 1 and 2 were determined by our CCE cutoff and the measured timeframe. If we changed the cutoff or modified the timeframe within our experiments, we would have changed our learner and non-learner populations as well. The separation of our subject population revealed a difference in directionality for mean RTs in Experiment 1—learners (based off of mean repeated-bound scene RTs – mean non-repeated scene RTs) were fastest in searching repeated-bound scenes, and were slowest in searching non-repeated scenes. Non-learners were fastest in searching repeated-unbound scenes, and were slowest in searching repeated-bound scenes. The learner vs. non-learner designation can reveal potential differences in learning performance, depending on the design of the experiment and the nature of the stimulus set.

3. Face vs. non-face search stimuli

Makovski (2017) demonstrated that for search items with distinct identities, the repeated binding of location and identity was required in order to facilitate search. The key feature of their stimuli is likely the semantic differences between search items. In order to examine whether or not the semantic differences between search items are necessary for search facilitation, we chose a stimulus set with high expertise but with low semantic differences between items. Our experiments' stimulus sets featured faces, in order to study learning for repeated contexts from stimuli with low semantic information, but for which participants have high viewer expertise. The stimulus set's faces shared the same ethnicity, the same facial expression, and similar facial features. We argue that most individuals have more experience viewing and processing faces compared against other search stimuli in the literature, such as letters (e.g. Chun & Jiang, 1998; Jiang & Chun, 2001; Conci et al., 2011; Zellin et al., 2013) and abstract contours (e.g. Endo & Takeda, 2004). All of the faces in our stimulus set were oriented upright, as most individuals were more likely to learn identity information from upright faces than inverted faces (Hochberg & Galper, 1967). Due to our stimulus sets' constraints, the faces appeared visually similar to one another, resulting in low stimulus uniqueness. We contrast this low stimulus uniqueness with Hout & Goldinger (2010) and Makovski (2016; 2017), where they featured images of real-world objects in their experiments belonging to multiple semantic categories, where versions of an object within a category appeared distinct from one another. Although faces are visually complex, we argue that our restrictive stimulus set's faces also hold little semantic information, compared against scenes with words (Goujon et al., 2007) and scenes with real-world objects (Hout & Goldinger, 2010; Makovski, 2016; Makovski, 2017). The constraints to our stimulus set reduced the likelihood that processing different semantic information between faces contributed to search facilitation (with the exception of the sex identification task in Experiment 2).

We expected to observe stronger contextual cueing effects for learning spatial information and identity-to-location binding information due to the greater viewer expertise. Learners in Experiment 1 had strong CCEs for repeated-bound scenes and for repeated-unbound scenes. This suggests that for learners, the presentation of both repeated contexts did not hinder their learning. Therefore, the repeated spatial information was likely weighted more than the repeated identity-to-location binding information. Otherwise, the learners would not have established a positive CCE for repeated-unbound scenes (where the identity-to-location binding was shuffled every block).

Learners in Experiment 2 had strong CCEs for repeated scenes with both same and different faces during the learning phase. At the start of the testing phase, we expected CC for different-faces scenes to decrease, and CC for same-faces scenes to remain the same, compared against the end of the learning phase. Instead, mean CCEs for both same-faces scenes and different-faces scenes decreased at the start of the testing phase. The break between phases may have contributed to the decrease in CCE for different-faces scenes, as it likely contributed to the decrease in CCE for same-faces scenes. Therefore, it is still possible that the learners weighted spatial information over identity-to-location binding for search facilitation.

The nature of an individual's learning of repeated contexts for a given stimulus type depends on the individual's viewer expertise for that stimulus type. Viewers did not learn repeated identity-to-location binding information from search arrays with abstract contours (Endo & Takeda, 2004). The low viewer expertise from the shape and artificial nature of the contours may have contributed to the lack of learned binding information. In contrast, Makovski (2016; 2017) observed learning for repeated identity-to-location binding information for search arrays with real-world objects. These objects were easily identifiable by subjects, suggesting that the subjects had high viewer expertise for real-world objects. The high viewer expertise may have facilitated the learning of binding information from the repeated scenes. In Zheng et al. (2019), binding information was learned for search arrays with faces. Similar to Makovski (2016; 2017), the high viewer expertise for faces may have facilitated the learning of binding information. But in our experiments, participants did not learn binding information, despite viewing search arrays with faces.

4. Future directions

The visual characteristics of our stimuli did not vary greatly from one another. Although faces are complex stimuli, our stimulus sets minimized possible identifying characteristics for any individual face, such that the stimulus sets featured low semantic information and low uniqueness between stimuli (except for Experiment 2, with one female target face out appearing with eleven male distractor faces). In both experiments, it is likely that spatial information learning dominated over identity-to-location binding learning. Viewers may have implicitly prioritized learning spatial information over identity-to-location binding due to the low semantic information between stimuli. If a stimulus set is rich in semantic information between stimuli (such as in Makovski (2016; 2017), with real-world objects), viewers may be more likely to learn specific identity information from each stimulus, which may contribute to learning repeated identity-to-location binding information. Varying the amount of displayed facial features, sex, facial expressions, and ethnicities can increase the semantic information and uniqueness between stimuli. For a stimulus set rich in semantic information, the proportions of learners and non-learners may change, due to individual differences in learning performance, and differences in individual viewing expertise.

5. Conclusion

Our experiments offer observations of learning when identity-to-location binding is present. However, learning for location information appears to be stronger than learning for identity-to-location binding. In Experiment 1, the similar magnitudes of mean CCEs for repeated-bound scenes and repeated-unbound scenes suggest that location information learning was weighted greater than identity-to-location binding. In Experiment 2, CC for same-faces scenes decreased following the shuffling of identity-to-location binding at the beginning of the testing phase, when we only expected CC for different-faces scenes to decrease. When a scene's identity-to-location binding information is repeated, spatial information and identity information are both repeated. But learning spatial information appears to be weighted more over learning identity information when the stimuli are visually similar to one another, and are low in semantic information. The visual similarities between the faces in our stimulus sets may have contributed to the weighting of learning spatial information over learning identity-to-location binding information, due to less learning for identity information. For a given stimulus set, greater visual differences between stimuli may be required to observe learning for identity-to-location binding information. The greater the uniqueness and semantic information for each stimulus, the greater the likelihood viewers would learn repeated identity information. If learning for spatial information is already strong, improved learning for identity information can promote learning for identity-to-location binding.

References

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 699–706.

Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, *22*(2), 153–176.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive psychology*, 36(1), 28–71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*(4), 360–365.

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 224–234.

Conci, M., & Müller, H. J. (2012). Contextual learning of multiple target locations in visual search. *Visual Cognition*, *20*(7), 746–770.

Conci, M., Sun, L., & Müller, H. J. (2011). Contextual remapping in visual search after predictable target-location changes. *Psychological Research*, *75*(4), 279–289.

Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Attention, Perception, & Psychophysics*, *66*(2), 293–302.

Glicksohn, A., & Cohen, A. (2011). The role of Gestalt grouping principles in visual statistical learning. *Attention, Perception, & Psychophysics*, *73*(3), 708–713.

Goujon, A., Didierjean, A., & Marmèche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition*, *15*(3), 257–275.

Goujon, A., Didierjean, A., & Marmèche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 50–71.

Goujon, A., Didierjean, A., & Poulet, S. (2014). The emergence of explicit knowledge from implicit learning. *Memory & cognition*, *42*(2), 225–236.

Hochberg, J., & Galper, R. E. (1967). Recognition of faces: I. An exploratory study. *Psychonomic Science*, *9*(12), 619–620.

Horstmann, G. (2007). Preattentive face processing: What do visual search experiments with schematic faces tell us?. *Visual Cognition*, *15*(7), 799–833.

Hout, M. C., & Goldinger, S. D. (2010). Learning in repeated visual search. *Attention, Perception, & Psychophysics*, *72*(5), 1267–1282.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology: Section A*, *54*(4), 1105–1124.

Kunar, M. A., & Wolfe, J. M. (2011). Target absent trials in configural contextual cuing. *Attention, Perception, & Psychophysics*, *73*(7), 2077.

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, *129*(4), 559–574.

Lleras, A., & Von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial vision*, *17*(4), 465–482.

Makovski, T. (2016). What is the context of contextual cueing? *Psychonomic bulletin & review*, *23*(6), 1982–1988.

Makovski, T. (2017). Learning "what" and "where" in visual search. *Japanese Psychological Research*, *59*(2), 133–143.

Makovski, T. (2018). Meaning in learning: Contextual cueing relies on objects' visual features and not on objects' meaning. *Memory & cognition*, *46*(1), 58–67.

Makovski, T., & Jiang, Y. V. (2010). Contextual cost: When a visual-search target is not where it should be. *Quarterly Journal of Experimental Psychology*, *63*(2), 216–225.

Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain*, *124*(7), 1417–1425.

Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception*, *17*(4), 475–481.

Vaskevich, A., & Luria, R. (2018). Adding statistical regularity results in a global slowdown in visual search. *Cognition*, *174*, 19–27.

Williams, M., Moss, S., Bradshaw, J., & Mattingley, J. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual cognition*, *12*(1), 29–50.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of experimental psychology*, *81*(1), 141–145.

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2013). Here today, gone tomorrow—adaptation to change in memory-guided visual search. *PloS one*, *8*(3), e59466.

Zellin, M., von Mühlenen, A., Müller, H. J., & Conci, M. (2013). Statistical learning in the past modulates contextual cueing in the future. *Journal of vision*, *13*(3), 19.

Zheng, L., Wang, Y., Wang, C., Jin, L., Bai, X., & Sun, H. (2019). The role of distractor identity in contextual cueing effect. Manuscript in preparation. Tianjin Normal University, China. McMaster University, Canada.