Differential contributions of global, local and background contexts in contextual-guided visual search

Xuelian Zang

SIGILL · UNIVERSITAT · LUDOVIC · MAXIM ·

München 2014

Differential contributions of global, local and background contexts in contextual-guided visual search

Inaugural-Dissertation
Zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität
München

vorgelegt von
Xuelian Zang
Aus Anhui, China

München, März 2014

Referent: Prof. Dr. Hermann J. Müller

Korreferent: PD. Dr. Zhuanghua Shi

Tag der mündlichen Prüfung: 02.Juli 2014

Table of Contents

Chapter 1. Introduction

1.1. Contextual-guided visual search

The Nobel peace laureate Albert Schweitzer once said, “Just as the wave cannot exist for itself, but is ever a part of the heaving surface of the ocean, so must I never live my life for itself, but always in the experience which is going on around me”. In our everyday activities, no object appears in isolation, but rather exists within a rich environment that contains both neighbouring objects and surrounding foreground/background contexts. To efficiently identify a target in an environment with an overwhelming amount of visual information, we have to select relevant while ignoring non-relevant information based on contextual analysis. For example, searching for a pedestrian in a picture would be more efficient if relevant regions of the picture (e.g., the roads) can be first quickly identified.

The ability to efficiently locate an object in a complex environment has been extensively studied over the last decades (e.g., Müller, Heller, & Ziegler, 1995b; Müller & Humphreys, 2003; Treisman, 1985; Treisman & Gormican, 1988; Wolfe, 2003b), and many factors concerning bottom-up and top-down processing of visual search have been identified: for instance, the bottom-up processing is driven by basic features and dimensions (e.g., color, motion, various depths cues, etc.), and the top-down processing is a goal-driven process that can be affected by various task requirements and prior knowledge (Wolfe, 1994a).

It is worth noting that, although filled with exhaustive visual information our visual world is often highly structured and stable over time. The learning and usage of these invariant relationships between different visual information - spatial context- can serve as an important factor to facilitate visual search (Chun, 2000). By using the spatial context of unchanged regularities in an environment, human and other animals can easily find an object placed in that environment. Take a basic task of searching for a book as an example, assuming a librarian knows well about the books in a library, it would be fairly straightforward for him/her to find a particular book that a user requested. As a regular worker in the library, the librarian has built a spatial map of the book order in his/her mind, thus helping him/her to quickly locate the book. But sometimes such helpful guidance may fail when the book is accidentally misplaced in a new location. In other words, the learned spatial regularities, such as the spatial relationship among bookshelf and surrounding books which are associated with the target (here the searched book), can be a useful spatial cue in helping the librarian's book search, when the spatial context remains unchanged.

The facilitation of invariant context in visual search, captured in the contextual cueing paradigm, has been investigated first by Chun and his colleague (Chun & Jiang, 1998) and then by a number of follow-up studies (e.g., Chun, 2000; Chun & Jiang, 1999; Conci, Müller, & von Mühlenen, 2013; Kunar, Watson, Cole, & Cox, 2013; Ogawa & Watanabe, 2010; Tseng & Lleras, 2013). In a standard contextual cueing paradigm (e.g., Chun & Jiang, 1998; Chun & Jiang, 1999), participants are asked to search for a target letter "T" and discriminate its orientation among a number of distractors "L"s (see Figure 1.1). Unbeknownst to the participants, half of the trials have the same spatial configurations among the search items that are repeated once per block during the experiments (hereafter we referred to it as "old" context), while the other half have randomly reconfigured new spatial configurations among the search items, and are never repeated during the experiments (hereafter referred to as "new" context). As participants performed the task, a general learning facilitation is often observed over trials for both old and new configurations, characterized by a progressive reduction of search time. More importantly, the target discrimination is usually quicker when it appeared in the invariant context than in the new context (Figure 1.1). The response facilitation induced by the repeated context is referred to as the "contextual cueing effect"

(Chun, 2000; Chun & Jiang, 1998; Chun & Jiang, 1999; Chun & Jiang, 2003; Chun & Nakayama, 2000). Interestingly, this kind of facilitation is not explicitly known to participants, in other words, participants learned the spatial regularities of the old contexts incidentally. But nevertheless, the implicit context can guide participants’ attention efficiently to the target locations as compared to the new contexts. Note the contextual cueing effect is not tied to individual identities, but rather to spatial configuration. In a further experiment (experiment 2), Chun and Jiang (1998) demonstrated that changing the characteristics of the individual items (e.g., from “므” to “5”) while maintaining spatial configuration among search items did not affect the learned context facilitation. This suggests that contextual cueing is established not by learning the features of the stimuli, but by learning the spatial context.

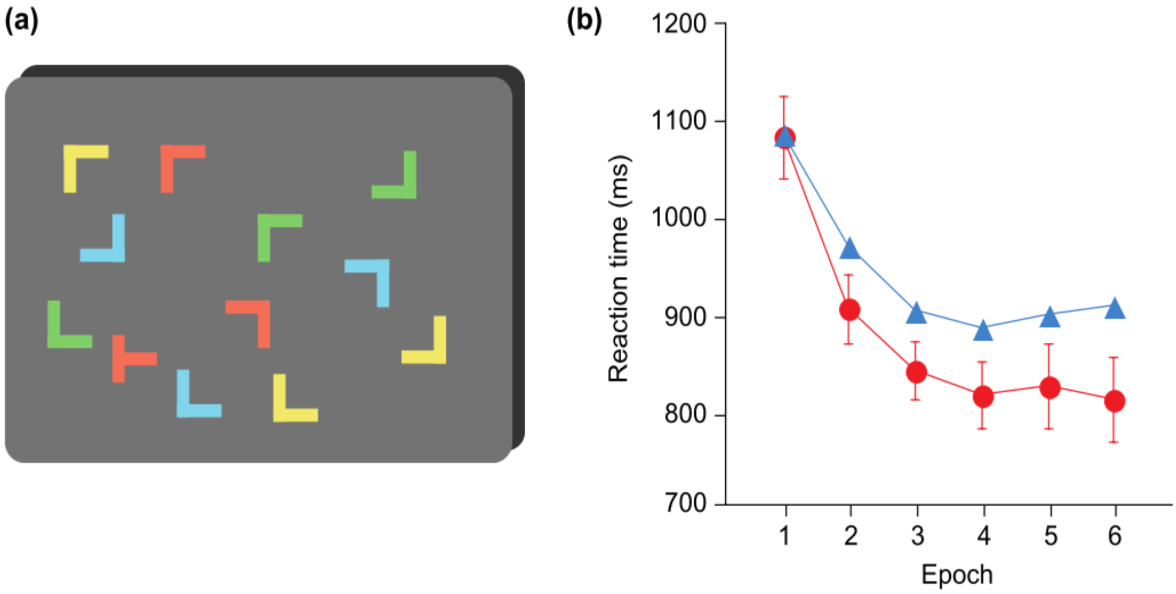

Figure 1.1 Schematic of spatial contextual cueing. a) A sample search array. The task was to search for a rotate T target amongst L distractors. b) Search performance as a function of epoch was faster for targets appearing in old configurations versus targets appearing in new configuration [This figure is reproduced from (Chun, 2000)].

Following the seminal work of Chun and Jiang (1998), an ample amount of studies have been conducted to better understand the underlying mechanisms of the contextual cueing effect. To name just a few, studies have investigated various factors that may influence the contextual facilitation in visual search, such as object and scene contexts (e.g., Rosenbaum & Jiang, 2013; Van Asselen, Sampaio, Pina, & Castelo-Branco, 2011), awareness of contextual learning (e.g., Schlagbauer, Müller, Michael, & Thomas, 2012; Smyth & Shanks, 2008), neural correlates (e.g., Manginelli, Baumgartner, & Pollmann, 2013; Westerberg, Miller, Reber, Cohen, & Paller, 2011), and eye movements behaviour during the contextual-guided visual search (e.g., Geringswald, Baumgartner, & Pollmann, 2012; Tseng & Li, 2004; Zhao et al., 2012). In particular, the effects of various types of contexts, such as the global structure, local spatial configuration, and background features, have been widely investigated (Brady & Chun, 2007; Brockmole, Castelhano, & Henderson, 2006; Brooks, Rasmussen, & Hollingworth, 2010; Geringswald et al., 2012; Kunar, Flusberg, & Wolfe, 2006; Kunar, John, & Sweetman, 2013; Olson & Chun, 2002; Song & Jiang, 2005; Van Asselen & Castelo-Branco, 2009). However, the findings of these studies are inconclusive, and sometimes contradict each other. For example, a recent study (e.g., Brockmole et al., 2006) found that the well-established contextual cueing was maintained when global context was kept constant (local context varied), but not when local context was kept constant (global context varied), thus suggesting that the learning of global context is more important for contextual cueing . On the contrary other studies found that local context is more important for the maintenance of contextual cueing (e.g., Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005). To date, however, the interaction between different contexts, such as global-local interaction, or background-foreground interaction has been rarely examined (see Section 1.2). Given that the local, global and background, foreground contexts usually coexist in a visual scene, it is crucial to explore their interactions for better understanding of the mechanisms supporting contextual cueing.

On this ground, the present thesis was designed to clarify the roles of different contexts, namely the local, global, background, and foreground contexts, as well as their interactions in contextual cueing. Specifically, the present thesis aims to answer the following questions: 1) Can the contextual learning and retrieval be solely based on the local foveal context without any aid of peripheral information? Is the peripheral information necessary in the contextual retrieval? 2) To what degree can the contextual cueing be transferred from the learned context to the novel display? 3) How do the local, global and foreground/background contexts interact with each other in the contextual learning?

In the following sections (Sections 1.2-1.5), I will provide a more elaborate summary of the effect of the global, local, and background contexts observed in the contextual cueing, summarize the related open issues concerning contextual cueing (Section 1.6), and presented the research topics of the current thesis (Section 1.7).

1.2. Roles of global versus local context in contextual-guided visual search

It is generally agreed that global and local visual information is processed differently (Brockmole et al., 2006; Hochstein & Ahissar, 2002; Murphy, Torralba, Eaton, & Freeman, 2006; Navon, 1977). Global information often accesses our consciousness quicker than the local information (Hochstein & Ahissar, 2002; Navon, 1977; Schyns & Aude, 1994). Take the classic Navon figure (Navon, 1977) for example, a large recognizable shape (e.g., a letter "H") that is composed of copies of a smaller different shape (e.g., letter "S"), demonstrates that information processing is prioritized from the global to local manner analogically expressed by Navon as "forest before trees". This is also true for scene interpretation

(Biederman, Mezzanotte, & Rabinowitz, 1982; Schyns & Aude, 1994; Torralba, Oliva, & Castelhano, 2006), in which scene is first identified by low spatial frequency blobs and multi-scale orientation filters, and then at high detail edges. Global feature, such as low spatial frequency blobs, can be efficiently detected in a very short exposure (Schyns & Aude, 1994), suggesting recognition process relies on coarse scene (blobs) information at the very first stage, but on fine (edges) information at later stages. One purpose of prioritizing global processing is to quickly compute a salient map of the scene, which can be used for bottom-up attentional guidance (Itti, Koch, & Niebur, 1998).

The global-to-local processing for scene recognition poses an interesting question regarding contextual-guided visual search: Are context learning and retrieval subject to a global-to-local processing? How does global- and local-context information contribute to the contextual cueing effect? While a number of studies have shown that the global invariant context is crucial to develop contextual cueing (e.g., Brockmole et al., 2006; Kunar et al., 2006), equal number of other studies have pointed out a dominant role of the local invariant context that plays a key role in contextual learning and retrieval (e.g., Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005). For example, Brockmole et al. (2006) demonstrated that participants tended to associate the global context to the target's location when they were asked to search for a target letter embedded in a computer rendered virtual scene, such as indoor library room. When the global predictive context (e.g., surrounded objects) was changed, the learned contextual facilitation disappeared. However, when the local background (e.g., the table on which the target is located) was altered, the learned contextual facilitation was not affected. Their findings thus support the global preference in the contextual learning. Similarly, Kunar et al. (2006) showed that global non-spatial attributes, such as background colour or line patterns, can be used as a predictive context for facilitating visual search. Note that the effectiveness of global predictive context in developing contextual cueing does not rule out the role of the local context. For example, Olson and Chun (2002) demonstrated that contextual cueing associated more strongly to the local context near the target than the context far from the target. Moreover, Brady and Chun (2007) showed that contextual cueing developed even when the repeated context was only a part of display, in such a way that the invariant context was limited to the target quadrant.

In summary, a number of studies have shown that both global and local contexts contribute to the contextual cueing effect. Arguably, however, owing to the fact that these studies mostly examined the influences of global and local context separately, their results reveal little about how local and global contexts interact during contextual learning and retrieval. Furthermore, in the studies that found global context to be a predominant factor for the contextual learning (Kunar et al., 2006; Rosenbaum & Jiang, 2013), the “global” was often referred to as the background colours or scenes that are separable from the search array items. By contrast, in those studies supporting the importance of local context (e.g., Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005), the term “global context” often meant the global structure of the search array, not separated from the local search items. As a result, the local and global contexts effects have not been neatly separated on the visual search stimuli of the current studies. In order to solve this problem, gaze-contingent technique was employed in one of my studies (see Chapter 2).

1.3. The transfer of contextual cueing after the configuration changes

As reviewed above, contextual cueing is a facilitation effect that mainly comes from the implicit learning of the invariant spatial context. And it has been shown that the learned context can be maintained for at least one week (Chun & Jiang, 2003; Jiang, Song, & Rigas, 2005). By contrast, the learned context is sensitive to changes of the spatial configuration, that is, it is relative inflexible to transfer the contextual cueing from an old display to a new display (with some variation from the old display). For example, convergent evidence has been gained that the learned contextual cueing diminishes when the target, and only the target, is re-positioned elsewhere in an old display (Chun & Jiang, 1998; Manginelli & Pollmann, 2009; Zellin, Conci, Von Mühlenen, & Müller, 2013; Zellin, Von Mühlenen, Müller, & Conci, 2014). Brady and Chun (2007) found that the contextual cueing disappeared when the old context at the target quadrant was moved to a different quadrant of the display. Similarly, Endo and Takeda (2005) found the learned contextual cueing could not be transferred when the context in sub-areas of the display was moved to a new location (e.g., swap the upper and lower half panel of the display while keeping the configurations in each half panel constant).

The inflexibility of the contextual cueing suggests that the learned context has a limited power of visual guidance. When some part of spatial context is changed, the cueing effect can be easily destroyed. Other studies have shown that one must preserve some predictability to maintain contextual cueing effective when the context is changed, and the degree of predictability is positively correlated with the magnitude of the transfer effect of contextual cueing (Chua & Chun, 2003; Conci & Müller, 2012; Conci, Sun, & Müller, 2011; Song & Jiang, 2005; Zellin, 2012). For example, Chua and Chun (2003) have shown that the magnitudes of the contextual cueing is negatively correlated with the degree of angular changes of the learned 3D spatial displays. A similar finding was observed by Makovski and Jiang (2010), in which the transfer of contextual cueing effect decreased as the target appeared further away from its original learned location. Jiang and Wagner (2004) provided similar evidence that the contextual cueing effect is not affected by the rescaled- or displaced-display, as the rescaled or displaced “old” display still preserves the predictive power. Those studies thus provide convergent evidence that predictability based on the invariant context is a key factor for maintaining contextual cueing when the learned context is altered to a certain degree.

1.4. Oculomotor behaviors in contextual-guided visual search

Eye movement is an important signature of visual information processing. For example, important features of a scene that are relevant to the on-going task are often fixated longer than others. Land, Mennie, and Rusted (1999) have shown that in a series of meaningful actions, such as making a cup of tea, almost all fixations that were made were directed to the object or objects involved in the current action (e.g., participants tend to fixate at the kettle during the action “find the kettle” or “lift the kettle”). Thus oculomotor data, in addition to manual responses, would help us to find the important information flow and the related mechanisms during cognitive search task, such as the development of contextual cueing during the visual search. Indeed, several studies of contextual cueing have shown a reduction of number of fixations during search of repeated displays. More interestingly, Peterson and Kramer (2001) revealed that in some trials, implicit memory was able to precisely guide saccade to the target location immediately after the onset of the repeated visual search display. Tseng and Li (2004) took a further look at a number of oculomotor parameters that might accompany the learned repeated displays. However, the only different oculomotor behaviour for the old display, compared to the new display, was the number of saccade and inter saccadic fixation duration. They also found that the visual search involves two phases: the initial ineffective and the subsequent effective search phases. During the initial ineffective search phase, the eye movements were not monotonically (or consecutively) getting closer to the target location, while in the effective search phase, eye movements are directly driven towards the target location. In another recent study, Manginelli and Pollmann (2009) examined oculomotor behaviour using a “misleading" contextual visual search task.

Participants were first trained with a standard contextual search display. A significant contextual cueing paired with a reduction of number of fixations for the repeated displays was found in the training session. Similar to the findings aforementioned, an effective search phase (characterized by earlier onset of a monotonic gaze approach phase towards the target location) was observed for repeated displays (e.g., Peterson & Kramer, 2001; Tseng & Li, 2004). In a subsequent test blocks, the target's location, but not the locations of distractors, was changed. With this non-predictable variation, the learned context could not cue the target's location anymore; instead, it could mislead participants' attention to the old target location, where the target was not there. The change of the target's location diminishes the contextual cueing effect. In addition, the reduction of the number of fixations and the effective search phase were gone during the test session (Manginelli & Pollmann, 2009). While the reduction of the number of fixations / saccades has been consistently observed in repeated displays, whether the repeated context affects the fixation duration is still unclear. For example, Van Asselen and Castelo-Branco (2009) found a reduction of the mean fixation duration for the old display compared to the new display. However, another recent study (Zhao et al, 2012) failed to support the reduction of inter-saccadic fixation duration.

In summary, a number of recent studies have provided deeper understanding about the contextual cueing mechanism by monitoring participants' eye movement behaviour. In general, it has been shown that the learning of repeated context leads to a reduction of the number of saccades, possibly because the repeated context, compared to the new context, guides participants' attention towards the target location. However, divergent results on the fixation duration in contextual cueing cry for further investigations.

1.5. Awareness in contextual learning - implicit vs. explicit learning

Implicit learning, although without a precise conceptual definition, usually refers to the learning of complex information in an incidental manner, without awareness of what has been learned. On the contrary, learning under hypothesis-driven and with fully conscious is consider as explicit. The differences between implicit and explicit learning have been summarized by Dienes and Berry (1997): 1) implicit, rather than explicit knowledge, is often relatively inflexible in transfer to different domains, 2) implicit, rather than explicit, learning occurs when attention is focused on specific items and not underlying rules, and 3) implicit learning and the resulting knowledge are often relatively robust.

To examine the awareness of contextual cueing, a typical recognition task, in which participants had to discriminate which display they had seen before, is usually adopted at the end of the contextual cueing search task (e.g., Chun & Jiang, 1998, 2003; Pollmann & Manginelli, 2009). The results often reveal that the old and new configurations cannot be distinguished (coded as implicit memory), due to contextual cueing been mainly driven by the incidental learning of old arrangements. While a great number of studies support implicit nature of contextual cueing (Jiang & Swallow, 2013; Jiang, Swallow, & Capistrano, 2014; Jiang, Won, & Swallow, 2014; Manginelli, Baumgartner, et al., 2013; Tseng & Lleras, 2013; Zellin, von Mühlenen, Müller, & Conci, 2013), other studies have questioned such claim (Brockmole & Henderson, 2006b; Conci & von Muhlenen, 2009; Geringswald et al., 2012; Geringswald, Herbik, Hoffmann, & Pollmann, 2013a; Geyer, Shi, & Müller, 2010; Rosenbaum & Jiang, 2013). For instance, when search for a target in a nature scene, which contains rich visual information, the learning of the scene is usually explicit (Brockmole & Henderson, 2006a, 2006b; Rosenbaum & Jiang, 2013). It's worth noting that the repeated old context can also be learned explicitly when searching for a target among distractors, characterized by significant (or sometimes marginally significant) higher hit rates compared to false alarm rates during recognition tasks (Conci & von Muhlenen, 2009; Geringswald et al., 2012; Geringswald et al., 2013a; Geyer, Shi, et al., 2010; Heeger, 1997; Macmillan, 2002; Shi, Zang, Jia, Geyer, & Müller, 2013). Using concurrent access-consciousness paradigm, Smyth and Shanks (2008) and Schlagbauer et al. (2012) have further suggested that some of the spatial configurations are accessible to awareness.

In short, contextual cueing learning is largely based on implicit long-term memory, but also supported in some degree by explicit memory. Explicit or implicit learning depends on different types of visual stimuli, such as nature scene or non-scene search arrays. When nature scenes are used as visual stimuli, explicit learning is more likely to be involved (e.g., Brockmole et al., 2006; Brockmole & Henderson, 2006a; Rosenbaum & Jiang, 2013). In contrast, when the display consists of non-scene items (e.g., with an array of letters or numbers), by and large implicit learning is dominated on contextual cueing.

1.6. Open questions related to contextual-guided visual search

As reviewed above, most of the previous studies investigated the role of either the local (a number of visual search items near the target location) or the global context (e.g., visual search items far away from the target's location), whereas the interaction between them was rarely examined. Furthermore, the effect of different associations between background and foreground contexts in contextual cueing effect is also largely neglected in the literature. Given that visual context in nature environment often involves different types of contexts (e.g., local/global context or foreground/background contexts), investigating the roles played by these contexts in contextual-guided visual search would be crucial for understanding mechanisms of contextual cueing.

1.6.1. Do local and global contexts affect differentially on context learning and retrieval?

A number of studies have found that preserving the local context near the target location, rather than the global configuration, is critical for developing a contextual cueing effect (Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005). On the contrary, using natural scene Brockmole et al. (2006) revealed that the global context (i.e., the whole view of the scene) plays a more important role than the local context (e.g., a table inside library with target on top of it) in developing the contextual cueing effect. Several subsequent studies (e.g., Kunar et al., 2006; Kunar, John, et al., 2013) have found that the global context (e.g., background colour) is able to boost contextual cueing. It is worth noting that the local and global contexts haven’t been strictly manipulated independently in those studies. Often, the local and global contexts are mixed together in the presentation of the whole visual search display. In such a whole view display, there is no way to keep global context constant while changing the local context. The natural link between the local and global contexts may lead to confounding observations in previous studies. This begs a number of intriguing questions: Is a pure local invariant context, without peripheral global information, sufficient to generate a contextual cueing? Or is it necessary to include a certain amount of global context for context learning and retrieval? Do the local and global contexts interact in contextual learning or/and contextual retrieval? Can the constant context be explicitly learned when only the local context near the fixation is presented at any given time during visual search?

1.6.2. To what degree can contextual cueing be transferred from the learned to novel displays?

The transferability is an important measure of the flexibility of contextual cueing, which can provide guidelines for developing user-friendly applications. For example, suppose we get familiar with the arrangement of icons, including the weather-app icon, on a particular display mode (e.g., landscape) in an iPad. As a result we can find the weather app without any effort; however, when the display mode is changed from the landscape to the portrait mode, spatial configuration among icons will be changed accordingly. Can we still efficiently locate the weather app? This is a typical example of the transfer effect of contextual cueing.

As reviewed in section 1.3, a number of studies have provided insightful evidence of the transfer effect of contextual cueing (Chun & Jiang, 1998; Makovski & Jiang, 2010; Manginelli & Pollmann, 2009; Zellin, 2012; Zellin, Conci, et al., 2013; Zellin et al., 2014). For instance, the transfer effect from the learned to novel displays is robust for geometric transformation that preserves the predictability of the target, but it can easily vanishes when the predictability is destroyed. However, none of the previous studies have investigated the transferability of contextual cueing when the display mode changes. Applying contextual cueing to the mobile application, we ask the following questions: is it possible to preserve the well-established contextual cueing when the display mode switching between landscape and portrait display mode? Are there any optimal remapping methods of icons rearrangements that can maximize the transferability of contextual cueing when shift the display modes? These open questions are examined in the second study of the present dissertation.

1.6.3. Does the foreground/background information of the search display affect contextual representation?

When the external visual world projects onto the retina, it forms a 2-dimensional retinotopic representation. In spite of this, we perceived a coherent world of meaningful objects that are effortlessly segregated from the background. This phenomenal experience arises from foreground-background segmentation processes (Caputo, 1996; Caputo & Casco, 1999). It has been shown that the segmentation processes substantially constrain attentional processes, as well as the reverse influence – the segmentation itself can be modulated by attention (Driver, Davis, Russell, Turatto, & Freeman, 2001b). For example, selective attention can push part of information to the background, thus it boosts relevant visual information processing, and facilitate search performance (Cave & Bichot, 1999; Wolfe, 2003b). Similarly, selective attention and foreground-background segmentation greatly influence the contextual representation. For example, it has been revealed that when the search display consisted of a white target "T" among black and white distractors ("L"s) in a contextual cueing task, only those distractors with the same colour as the target were constructed into contextual memory (Jiang & Leung, 2005). The distractors with target-unrelated colour were simply pushed to the background, and ignored. It should be noted however, the feature of foreground/background is not fixed, but rather depends on the tasks. For example, when you search trees on Google earth map, roads may consider as irrelevant background. However, when you search cars on the same map, roads could be a useful background, as the background (roads) and foreground (cars) are strongly coupled. Considering selective attention as an argument, several questions related to the role of foreground/background context in contextual cueing are still open: Would different item-independent visual information (e.g., foreground vs. background information) be encoded together with the foreground item configuration during spatial context learning and retrieval? Does contextual representation depends on the properties (i.e., foreground vs. background) of the visual information on the display? Does image segmentation interact to contextual cueing? These questions are examined in the third study of the present thesis.

1.7. Cumulative research work

To address those open issues stated above, the present Ph.D work mainly focus on the following research topics:

1.7.1. Learning locally, retrieving globally - Evidence of contextual cueing with gaze-contingent visual search task

The first study of the thesis (Experiments 1-3) was proposed to answer the question concerning interactions between the local and global contexts. In order to examine whether pure local information, in the absence of any global-structure information, is sufficient for contextual guidance in visual search during the training session of the present experiments, a gaze-contingent technique (e.g., Loschky & McConkie, 2000) was employed in the standard contextual cueing paradigm (e.g., Chun & Jiang, 1998), such that the local foveal information can be presented separately from the global peripheral information. In the following transfer sessions, the gaze-contingent was removed to examine if the availability of the peripheral global information helps contextual-guided visual search.

In the first experiment, the size of gaze-contingent view area (i.e., the visible region near the fixation) was set to 8°. On average, 2.09 items out of the total 12 items are visible. The results showed no contextual cueing effect in the training session, but significant facilitation with faster RT for repeated display compared to novel display in the transfer session, in which the whole display was visible. Moreover, the context facilitation was already visible during the first block of the free-view transfer session. The finding of Experiment 1 suggests that the repeated context can be learned with limited local information, but cannot be effectively retrieved in the absence of the peripheral information. One question remains, whether the local information is too scarce, such that the contextual cues are hard to retrieve. Thus, in a second experiment we extended the local visual area from 8° to 12°, on average, 4.69 items were visible inside view area (the number of items was doubled). This time we found a significant contextual facilitation in both training and transfer sessions, suggesting that information within the 12° view angle provides sufficient spatial configuration for both contextual learning and retrieval. Experiment 3 was designed to further investigated whether a brief preview (150 ms) of the global context prior to the gaze-contingent search display (8°) can aid contextual cueing retrieval. Again, Experiment 3 revealed significant contextual cueing effect in both sessions, in line with Experiment 2. Further eye movements analyse revealed that the contextual cueing facilitation was associated with reduced number of saccades and extended fixation duration for old display compared to new display.

In summary, the Study 1 investigated the interaction of local/global context in contextual guided search by employing gaze-contingent techniques. We found that repeated spatial context can be implicitly learned but not retrieved, based on scarce local information with only 2~3 visual items available under gaze-contingent limited view. In order to effectively retrieve the learned contextual cueing, some global peripheral information (e.g., the global brief preview or larger size of gaze-contingent view) must be available.

1.7.2. Transferability of contextual cueing in full-icon display remapping

The second study was designed to investigate whether learned contextual cueing can be transferred when display orientation (or display mode) was changed. Changes of the display modes (e.g., from the landscape to the portrait) happen regularly when you use a mobile device, such as an iPad. After a change of display mode, the icons on the display are shuffled: the positions of icons in one mode are remapped to the other mode by keeping the positional order (left to right and up to down) constant across all icons. Although this remapping method preserves the positional order and most of the horizontal inter-icon relationships, it destroys almost all local icon relationships (or local context), when the display is arranged as a rectangle (see more details in chapter 3). As we reviewed above, the transferability of the learned contextual facilitation will also be destroyed. To develop better icon-remapping methods for future mobile interaction, four experiments with “full-icon” displays were designed in the second study to investigate the transferability of the contextual cueing between two different display modes. Thus, besides the available “position-order” remapping method, three other remapping methods, namely, “global rotation”, “local invariant” and “central invariant” remapping methods were examined in four separate experiments. For the “global rotation” remapping method, the whole display rotated 90° clockwise when the display mode varied from landscape to portrait by keeping icon-icon spatial configuration constant. For the “local invariant” method, 5 local regions were kept constant after the change of the display mode, while the “central invariant” remapping methods kept the icons in the central maximal square region unchanged (see details in chapter 3). Each experiment includes three sessions: training, transfer and recognition sessions. The landscape displays were used in the training session, and in the subsequent transfer session, the same displays were remapped to portrait display mode according to the four different remapping methods. Last but not the least is the recognition session that is aimed to test whether participants had learned the context explicitly or not.

All experiments resulted in robust contextual cueing effects after the training. Interestingly, the learned context was only successfully transferred to the novel portrait mode for the “local invariant” or “central invariant” remapping methods, but not for the “position-order” and the “global rotation” remapping methods. The results suggest the traditional “position-order” remapping used in current mobiles is not optimal in helping user's search performance for display mode changes. Moreover, the “global rotation” remapping method, although it happens frequently during everyday life, does not facilitate users’ performance, partly due to requiring additional mental resource for mental rotation (Böckler, Knoblich, & Sebanz, 2011; Borst, Kievit, Thompson, & Kosslyn, 2011; Ionta & Blanke, 2009; Shepard & Metzler, 1971). Most importantly, the learned contextual cueing was preserved by keeping invariant local context at maximum across display modes using the “local invariant” or “central invariant” remapping methods. A further interesting finding of the second study is that more than 80% of participants reported that they have noticed the repetition of displays during the visual search task, suggesting that observers were able to explicitly recognize the displays, and remapping did not hamper explicit recognition.

1.7.3. Interaction between foreground/background item-independent information and configural context in contextual cueing

Our first study found that contextual cueing can be learned based on scarce local information; however, to effectively retrieve the learned context, a certain amount of visual information or global context must be available. In the second study, we found that the learned context can be transferred to a novel display mode that most of the local configurations were preserved. In the third study of three experiments (Experiment 8, 9 & 10), we focus on whether item-independent information can be encoded together with the spatial configural context, and whether the learning and expression of contextual cueing depends on the characteristics of the item-independent features (i.e., foreground or background information) on the visual search display. The visual search stimuli used in the present experiments was made up of a item-independent geometric shape (i.e., a drawing pseudo cuboid shape which is presented on 2D plane) and a standard foreground visual search array (i.e., a target “T” among a number of distractors “L”s) that were used as in the standard contextual cueing paradigm (Chun & Jiang, 1998). During the training session, the feature of the item-independent cuboid was manipulated as foreground in Experiment 8 but as background in Experiments 9 and 10. In a subsequent transfer session, the item-independent cuboid was either titled 90° or removed, to examine whether the learned contextual cueing can be transferred or not. The results showed a significant contextual cueing effect during training session in all the three experiments; however, the established cueing was transferred from the old to the novel display only when the cuboid was presented as background information (Experiment 9 and 10), but not when it was controlled as foreground information (Experiment 8). The findings suggest that the contextual cueing effect developed with the reference to both foreground item-independent shape and visual search items, whereas the cueing was not preserved when the foreground shape was varied or removed. By contrast, when the cuboid was presented as background, the contextual cueing developed solely based on the foreground configural information, rendering itself less vulnerable when the cuboid is changed. In conclusion, the involvement of item-independent shape during contextual cueing learning and retrieval depends on its foreground/background features on the visual search display.

1.8. Conclusion

In summary, the present thesis comprises of three studies that investigated the differential roles played by global-local, and foreground-background contexts, as well as the interactions among different contexts in contextual-guided visual search. By combining gaze-contingent technique with a standard contextual cueing paradigm in Study 1, we found that contextual cueing can be learned based on pure local information, but peripheral global information is needed in order to retrieve the learned contextual cueing. In the second study, we found that keeping the local context invariant is crucial to maintain the contextual cueing after the display mode varied (e.g., from landscape to portrait). In the third study, the feature (i.e., foreground vs. background) of the context was found to be an important factor in contextual learning and its retrieval: the foreground context is likely to be encoded while the background context is most likely to be ignored during the contextual learning.

Chapter 2. Invariant spatial context is learned but not retrieved in gaze-contingent limited-viewing search

2.1. Abstract

Our visual brain is remarkable in extracting invariant properties from the noisy environment, guiding selection of where to look and what to identify. However, how the brain achieves this is still poorly understood. Here we explore interactions of local context and global structure in the long-term learning and retrieval of invariant display properties. Participants searched for a target among distractors, without knowing that some, “old” configurations were presented repeatedly (randomly inserted amongst “new” configurations). Crucially, we simulated tunnel vision, limiting the visible region around fixation. Robust facilitation of performance for “old-” vs. “new” contexts was observed when the visible region was large, but not when it was small. However, once the display was made fully visible during the subsequent transfer phase, facilitation did become manifest. Furthermore, when participants were given a brief preview of the total display layout prior to tunnel view search with only two items visible, facilitation was obtained already during the learning phase. The eye movement results revealed contextual facilitation to be coupled with changes of saccadic planning, characterized by slightly extended gaze durations but a reduced number of fixations and shortened scanpaths for “old” displays. Taken together, our findings show that invariant spatial display properties can be acquired based on scarce, para-/foveal information, while their effective retrieval for search-guidance requires the availability (even if brief) of a certain extent of peripheral information.

Keywords contextual cueing, learning, memory retrieval, eye movements

2.2. Introduction

The ability to learn spatial context is vitally important for humans and other animals, as spatial context can improve the efficiency of foraging and other search tasks. In our everyday lives, we frequently use contextual cues to find a specific target object, for example, when looking for the stapler first in its “usual” place on the desk, when looking for the neuroscience book straight on the left side of the lower bookcase shelf as it is normally placed there, or when tapping immediately the weather icon at the center of the iPad display without much search effort as the task has been repeated many times. However, when such target objects are accidently misplaced to a “new” location, for instance, when the weather icon is shuffled to another position due to a change of the display mode (Shi et al., 2013), we often need additional time and effort to find them. The fact that spatial contextual information facilitates visual search, referred as to “contextual cueing”, has attracted much attention in recent years (for a review, see Chun, 2000; Oliva & Torralba, 2007).

Concerning spatial information, researchers generally agree that two types of information, namely: global and local spatial configuration, contribute to object localization and identification processes (Hochstein & Ahissar, 2002; Navon, 1977; Schyns & Aude, 1994). One still open question with regard to contextual cueing is how global and local context information interact during contextual learning and retrieval (Brady & Chun, 2007; Brockmole et al., 2006; Brooks et al., 2010; Kunar et al., 2006; Olson & Chun, 2002; Rosenbaum & Jiang, 2013; Song & Jiang, 2005). Global context has been shown to be an important factor in contextual learning and transfer (Brockmole et al., 2006; Brooks et al., 2010; Kunar et al., 2006; Rosenbaum & Jiang, 2013). For example, Brockmole et al. (2006) demonstrated that in search of naturalistic scenes, contextual cueing is biased to global-context associations. In their experiments, participants were asked to search and identify an arbitrarily located target letter within a computer-rendered realistic scene, where the “local” context was defined as a set of objects near the target, while the remainder of the scene was referred to as “global” context. Brockmole et al. (2006) found contextual cueing to manifest after repeated exposure to certain displays. However, the acquired cueing effects transferred from the learning to the test session only for search displays that maintained the global information, but not for displays that only maintained the local set of objects near the target. Similarly, Kunar et al. (2006) found predictive global background context (e.g., background scene colors or line patterns) to facilitate visual search. More recently, Rosenbaum and Jiang (2013) initially trained participants on displays in which the target locations were predicted by both background scene context and array-based context (i.e., the arrangement of the display items), and then tested participants either with displays that included the same background context but varied the array context, or displays that included the same array context but varied the background context. Rosenbaum and Jiang (2013) found that contextual facilitation was transferred from the training to the test phase when the background scene was the sole predictive search cue, but not with the array of search items as the sole cue – suggesting that global background scene contextual cueing precluded item-based cueing when both were predictive of the target location. However, in other studies without any manipulation of background context, local invariant array-based context appeared to be sufficient for generating contextual cueing - that is, cueing manifested even if (some of the) items beyond the local region underwent positional changes (Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005). For instance, Olson and Chun (2002) examined contextual cueing effects under partly invariant display configurations - in which the items in one half of the display were kept constant while the items in the other half varied in their positions. They observed a significant contextual cueing effect when the target appeared in the invariant, but not when it appeared in the other, half of the display. The importance of local context has also been confirmed by several other studies (Brady & Chun, 2007; Song & Jiang, 2005), all suggesting that partial local predictive information provided by repeated items near the target is sufficient to induce contextual cueing.

The studies reviewed above make valuable contributions to understanding the fundamental roles of global and local information in contextual learning and retrieval, by showing that both global and local context information can contribute to the cueing effects. However, global and local context are not explicitly defined in these studies. Local context often implicitly refers to the local spatial configuration near the search target (e.g., Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005), while the global context generally refers to background colors or scene properties, separable from the search items (e.g., Kunar et al., 2006; Rosenbaum & Jiang, 2013). This kind of global and local context information differs from eye-centered para-/foveal local and peripheral global information. In a typical search task, the eyes move from one location to another, continuously bringing display regions of interest into the fovea, guided by a map of overall-saliency or “priority” signals (the latter combining both bottom-up and top-down information) (Itti & Koch, 2001; Müller, Heller, & Ziegler, 1995a; Wolfe, 1994b). Recently, it has been shown that loss or degradation of foveal vision can eliminate the contextual advantage conferred by repeated displays in visual search (Geringswald et al., 2012; Geringswald et al., 2013a), suggesting that foveal local information plays a critical role in contextual learning and retrieval. On the other hand, degenerative eye diseases, such as retinitis pigmentosa (RP), often cause the loss of peripheral vision, resulting in a constricted, “tunnel” vision (Hartong, Berson, & Dryja, 2006). Yet, it remains unclear whether pure eye-centered local information, in the absence of any peripheral global-structure information, would be sufficient for contextual guidance in visual search.

The present study, of three experiments, was designed to fill this gap in our knowledge. If contextual cueing were solely based on para-/foveal local learning, the availability of local context should suffice for a contextual-cueing effect to manifest using simulated tunnel vision, in which only the items in the vicinity of each fixation position are visible. This hypothesis derives partially from several previous studies (Brady & Chun, 2007; Olson & Chun, 2002) suggesting that two to three local items near the target can provide enough information for contextual cueing to develop. By contrast, if the presence of global peripheral structure is necessary for learned contexts to be retrieved and guide visual search, then the lack of peripheral information should impede contextual cueing. To examine these alternative predictions, we simulated tunnel vision using a gaze-contingent viewing technique in a classic contextual-cueing visual search paradigm. Based on real-time tracking of the eye position, gaze-contingent tunnel viewing of the search display provides detailed local information within the central para-/foveal area, and only coarse information in the periphery (Loschky & McConkie, 2002; Loschky & McConkie, 2000; Parkhurst, Culurciello, & Niebur, 2000). Accordingly, in Experiments 1 and 2, we manipulated the size of the viewing tunnel to ascertain whether pure eye-centered local information is sufficient to generate a contextual cueing effect, as well as examining for differential oculomotor scanning behavior between old (repeated) and new display configurations. In Experiment 3, we further examined for potential cueing benefits deriving from brief previews of the global item layout for the subsequently performed gaze-contingent search of the target display.

2.3. Experiment 1

Experiment 1 examined whether para-/foveal local spatial context, in the absence of peripheral global structure, would suffice to engender a classic contextual-cueing effect, and whether the learned context could be transferred to the search display with the global configuration (i.e., whole display) presented without limitation. To this end, a standard contextual-cueing search paradigm (Annac et al., 2013; Chun & Jiang, 1998; Schlagbauer et al., 2012) was adopted in Experiment 1: participants were presented with a sequence of trials on which they searched for a target letter “T” among distractor letters “L”, and responded to the target T’s orientation. The experiment consisted of three sessions: a “training”, a “transfer”, and a “recognition” session. During training, the search display was visible only within a region around the gaze location at any given time (i.e., tunnel view). To examine whether the context acquired during the training session, if any, could be transferred to a free-view search condition, the whole display was visible in the transfer session. The final recognition session examined whether or not participants were able to explicitly tell apart repeatedly encountered (i.e., “old-context”) from ad-hoc generated (i.e., “new-context”) displays.

2.3.1. Materials and Methods

Participants. 13 participants (7 females, mean age: 23.8 years) with normal or corrected-to-normal visual acuity took part in Experiment 1. They gave written informed consent in accordance with the declaration of Helsinki 2008, and were paid for their participation. None of them were aware of the purpose of the study.

Apparatus. The experiment was conducted in a dark cabin (0.35 cd/m^2). The search display was presented on a 19-inch CRT monitor, with a refresh rate of 100 Hz, at a viewing distance fixed to 54 cm with the support of a chin rest. Movements of participants’ dominant eye were monitored using an Eyelink 1000 desktop-mounted system (SR Research Ltd., Canada), set at a sampling rate of 1 kHz. Stimulus presentation, response recording, and eye movement sampling were controlled via a Matlab program using the Psychtoolbox and the Eyelink Toolbox (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002).

Stimuli. The search display consisted of one “T”-shaped target and eleven “L”-shaped distractors (43.2 ±1.9 cd/m2). The stimuli were positioned randomly at 12 of the 44 possible locations within a circular display matrix, with a diameter of 16° of visual angle (see Figure

2.1A). Both “T” and “L” shapes were composed of two equal-length lines (1°), one horizontal and one vertical. In “T” stimuli, the lines’ contact point was at the tip of the vertical line and, respectively, the center point of the horizontal line; in “L”-shaped stimuli, the contact point was at the tip of the vertical line and the left side of the horizontal line with a 0.2° offset to the tip; the offset construction of the “L” stimuli was meant to increase the difficulty of the search task. In the search display, the L shapes could appear in one of the four orthogonal rotations, while the T shapes were rotated either 90° to the left or 90° to the right, requiring a “left” or, respectively, “right” response (see below).

Two types of the search configuration were constructed and presented in Experiment 1 (as well as the subsequent experiments): “old” and “new” configurations. “Old” configuration consisted of 8 randomly generated displays, which were kept unchanged during the whole experiment, and were presented on randomly selected trials within each block and repeated throughout the experiment. “New” configurations, by contrast, consisted of 8 newly generated displays for each block. To balance the target locations between “old” and “new” displays, target locations were distributed equally across the display’s 4 quadrants for both types of display; that is, for both “new” and “old” configurations, there were two randomly selected target locations in each quadrant (the only constraint being that the center four locations never contained a target). In addition, target orientation was randomized across the generated displays.

Participants viewed the (visible parts of) search display through a gaze-contingent, tunnel window. The location of the window depended on the participant’s current gaze coordinates, which were on-line updated through the Eyelink toolbox (Cornelissen et al., 2002). The default psychophysical sample configuration of the Eyelink 1000 (i.e., saccade velocity threshold set as 22°/s, saccade acceleration threshold set as $4000°/s^2$) was adopted for the eye data samples. Average delays from eye movement to position data availability were less than 10 ms, that is, the display was updated either immediately in the next refresh cycle or within a maximum delay of 20 ms (for eye movements that were detected only towards the end of a given screen refresh cycle). The visible tunnel was 8° in diameter and consisted of a fully visible central para-/foveal area of 5°, and an outer ring (from 5°- 8°) with a gradual transition, realized using a Gaussian blob filter, from fully visible to nonvisible information. The remainder of display was completely blank, providing no global information as to the layout of the search display (see Figure 2.1 B). On average, 2.09 out of the 12 display items were visible inside the gaze-contingent viewing area. When the gaze coordinates were unavailable due to eye blinks or signal losses, the display was kept completely blank.

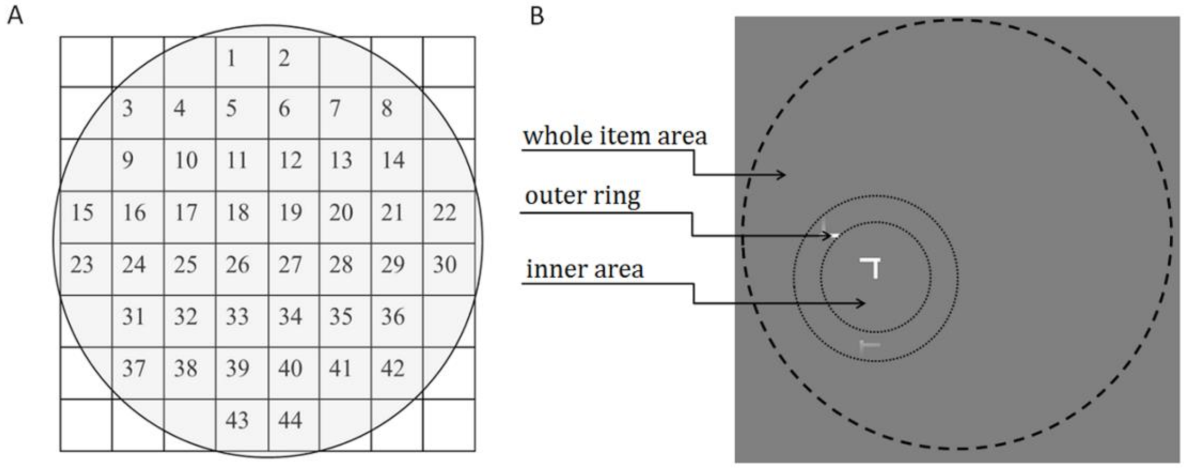

Figure 2.1 A) A spatial matrix with 44 possible locations was used in the experiments. Search items were randomly distributed across these locations. The grid, numbers, and the circle were invisible during the actual experiments. B) Example of a gaze-contingent search display with 8° in diameter. The items in the inner, central area around the current fixation position (5° in diameter) was fully transparent, while the items in the outer ring (between 5° and 8°) underwent a gradual transition from fully transparent to fully opaque using a Gaussian filter. The circles and arrows are drawn only for illustration.

Design and Procedure. The experiment consisted of three sessions: training (25 blocks), transfer (5 blocks), and recognition (1 block). Each block consisted of 16 trials, with 8 “old” and 8 “new” configurations. All visual search displays were presented under conditions of dynamic tunnel viewing in the training session, while the whole displays were visible during the transfer and the recognition session.

In the training and transfer sessions, participants were asked to discriminate the orientation of a target letter “T” (randomly oriented either 90° or 270°) among distractor “L”s (randomly oriented 0°, 90°, 180° or 270°) as fast and accurately as possible by pressing a key, either the left or the right arrow key on the keyboard, using their index fingers. A trial started with the appearance of a central fixation point which participants were instructed to fixate. The fixation marker disappeared after 500 ms of continuous gaze, immediately followed by a search display. Participants were allowed to search the display freely without any restriction on making eye movements. The search display disappeared when a response was made or when the presentation exceeded 15 seconds. After a random interval of 1.0–1.2 seconds, the next trial started (see Figure 2.2).

The last, recognition session included the original 12 repeated and another 12 newly generated displays, with both types of display presented in free view and in randomized order. Participants were asked to indicate whether or not they had previously seen a given display during the course of the experiment. The display was presented on the screen until the response was made. Participants were expressly told that about half of the displays were repeated and the other half new. No feedback about the correctness of the answer was given.

To increase the power of statistical analyses, every 5 consecutive blocks were grouped into epochs, forming the epochs 1–5 for the training session and epoch 6 for the transfer session.

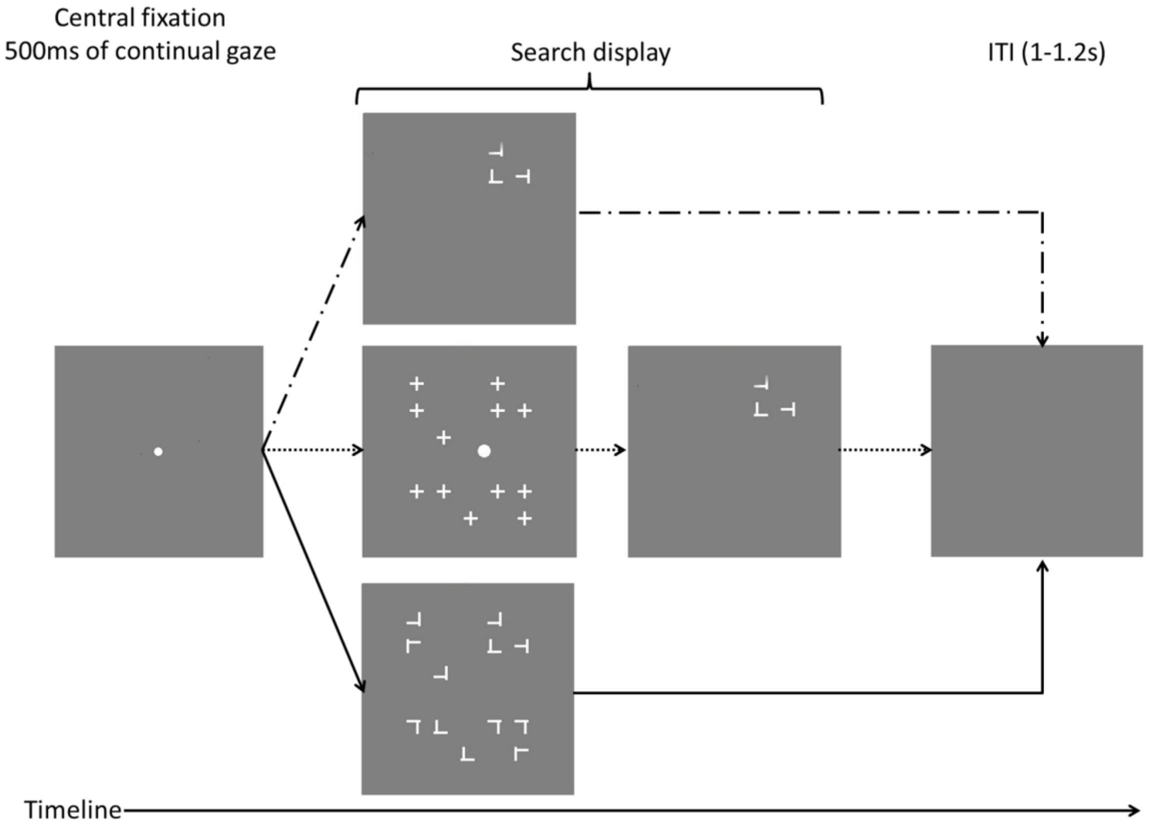

Figure 2.2 Schematic illustration of the search paradigm. The upper path (marked by the dash-dot lines) indicates the procedure of gaze-contingent limited-viewing search implemented in the training phases of Experiments 1 and 2. The middle path (marked by the dashed lines) indicates the procedure of (brief, 150ms) preview display followed by gaze-contingent limited-viewing search implemented in the training phase of Experiment 3, and the lower path (marked by the solid lines) illustrates the procedure in the transfer phase of all three experiments. ITI = inter-trial interval.

Data analysis. Invalid gaze samples due to eye blinks or signal losses were discarded. Furthermore, nearby short fixations (< 50 ms) separated by small movement distances (< 0.5°) were classified as gaze “dwells” and merged into a single gaze episode, using a custom-made Matlab script. Following these refinements, the average fixation duration was 273 ms.

Three different measures of oculomotor behavior were calculated for each trial: (i) the number of fixations during the search, (ii) the mean fixation duration, and (iii) the efficiency of the scanpath towards the target. For measures (i) and (ii), the first fixation was not included to avoid possible contamination by the initial mandatory fixation. Measure (iii) was operationalized by calculating the “scanpath deviation”, that is, the difference between the total saccadic distance of the fixations during the actual scanpath (i.e., the total length of the actual scanpath) and shortest distance from initial fixation to the target (i.e., the length of the shortest possible scanpath). This measure is similar to the “scan pattern ratio” (i.e., the ratio between the total distance covered by the eye during the search and the shortest scanpath) used in previous studies (Brockmole & Henderson, 2006a; Geringswald et al., 2012; Geringswald et al., 2013a). Arguably, the “scanpath deviation” measure is preferable because it avoids distortions introduced by extreme denominators in the “scan pattern ratios”, thus ensuring normally distributed data for further statistical analysis.

Repeated-measures analyses of variance (ANOVAs) were carried out on reaction times (RTs), number of fixations, fixation durations, and scanpath deviations, with degrees of freedom Greenhouse-Geisser corrected if the sphericity assumption was violated. Further LSD contrast tests were carried out as necessary. In addition, JZS Bayes factors (null/alternative) (Rouder, Speckman, Sun, Morey, & Iverson, 2009) were calculated for those results that favoured the null hypothesis. According to Jeffries (1961), a value greater than 3 provides “substantial” evidence for choosing the null hypothesis.

2.3.2. Results

Trials with erroneous responses or reaction times (RTs) outside the range 0.2 s and 10 s were excluded from analysis. Both the overall mean error and outlier rates were low (errors: 1.15%; outliers: 3.93%). The error rates were comparable across all conditions: context, $F(1, 12) = 3.25$, $p = 0.10$, $\eta_p^2 = 0.21$; epoch, $F(2.43, 29.14) = 2.2$, $p = 0.12$, $\eta_p^2 = 0.16$; and interaction, $F(5,60) = 0.66$, $p = 0.66$, $\eta_p^2 = 0.05$, revealing no evidence of performance accuracy improving as a result of training.

***Training session*.** Two-way repeated-measures ANOVAs, with context (old vs. new) and epoch (1-5) as factors, were applied to the (individual-condition) mean RTs, the mean number of fixations, the mean scanpath deviation, and the mean fixation duration. Note that the major part of the RT is the time required for the search (as compared to the time taken for discerning the target orientation and selecting and executing the appropriate response), and the search time can be "reconstructed" from the number of fixations and the respective fixation durations. Accordingly, some similar result patterns were expected among these analyses (see Figure 2.3).

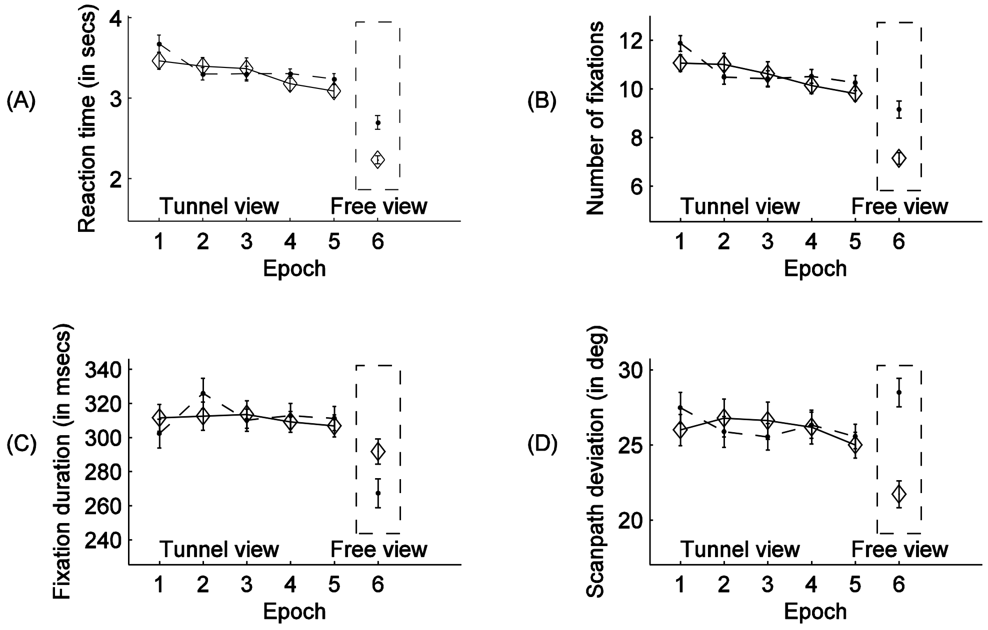

Figure 2.3. Results of Experiment 1. Mean reaction times (A), mean number of fixations (B), mean fixation duration (C), and scanpath deviation (D), all with associated standard errors, are shown as a function of experimental epoch and display context ("old", indicated by solid-diamond lines, vs. "new", indicated by dash-dot lines). Epochs 1-5 represent the training session with the tunnel view display, and Epoch 6 (in the dashed box) represent the transfer session with the free-view display.

A significant effect of procedural learning (over epochs) was observed in both the RTs and the number of fixations, with an average reduction of 406 ms in RT, $F(4,48) = 5.03$, $p < 0.01$, $\eta_p^2 = 0.30$, and 1.44 fewer fixations, $F(4,48) = 5.84$, $p < 0.001$, $\eta_p^2 = 0.33$, in the epoch 5 compared to epoch 1. However, there were no significant changes over epochs in the scanpath deviation, $F(4,48) = 0.61$, $p = 0.66$, $\eta_p^2 = 0.05$, and the mean fixation duration, $F(4,48) = 0.93$, $p = 0.46$, $\eta_p^2 = 0.07$.

Of major importance, in the training session, we failed to observe any contextual-cueing effect, that is, any difference between "old" and "new" configurations, in any of the four measures: mean RTs, $F(1,12) = 0.35$, $p = 0.57$, $\eta_p^2 = 0.03$, JZS Bayes factor (null/alternative) = 2.60; mean number of fixations, $F(1,12) = 0.16$, $p = 0.70$, $\eta_p^2 = 0.01$, JZS Bayes factor = 3.16; scanpath deviation, $F(1,12) = 0.001$, $p = 0.98$, $\eta_p^2 = 0.001$, JZS Bayes factor = 4.22; and mean fixation duration, $F(1,12) = 0.04$, $p = 0.85$, $\eta_p^2 =0.003$, JZS Bayes factor =4.72. The Bayes factors were close to or greater than 3, indicative of "some" to "substantial" evidence in favor of the absence of a contextual-cueing effect in the training session. In addition, the two-way (epoch × context) interactions were non-significant in all four measures: RT, $F(4,48) = 1.72$, $p = 0.16$, $\eta_p^2 = 0.13$; mean number of fixations, $F(4,48) = 1.75$, $p = 0.15$, $\eta_p^2 = 0.13$; scanpath deviation, $F(4,48) = 1.13$, p = 0.35, $\eta_p^2 = 0.09$; mean fixation duration, $F(4,48) = 0.82$, $p = 0.52$, $\eta_p^2 = 0.06$. These results indicate that gaze-contingent tunnel viewing - providing only limited, local information during search - prevents either contextual learning altogether or the retrieval of learned context.

Transfer session. Interestingly, when comparing "old" to "new" contexts in the transfer session (epoch 6), the contextual-cueing effect turned out significant and manifest in all four measures: there was a reduction of 466 ms in mean RT, $F(1,12) = 23.89$, $p < 0.001$, $\eta_p^2 = 0.67$; a reduction of 2.01 in the mean number of fixations, $F(1,12) = 23.93$, $p < 0.001$, η_p^2

= 0.67; a reduction of 6.76° in scanpath deviation, $F(1,12) = 35.60$, $p < 0.001$, $\eta_p^2 = 0.75$; and an increase of 24.41 ms in the mean fixation duration, $F(1,12) = 8.38$, $p < 0.05$, $\eta_p^2 = 0.41$. A further one-way repeated-measures ANOVA of the mean RTs in the very first block (of the 5-block epoch) of the transfer session revealed significantly faster responses – a very substantial reduction of 719 ms – for “old” compared to “new” contexts, $F(1,12) = 35.36$, $p<0.01$, $\eta_p^2 = 0.75$, strongly suggesting that the contextual-cueing effect, rather than being due to fast learning in the transfer session, was transferred from the training session.

To examine more closely whether dynamic tunnel viewing causes a change of oculomotor and search behavior, epoch 5 (last epoch of tunnel viewing) and epoch 6 (free viewing) were compared using two-way ANOVAs. The results revealed a substantial drop of mean RT in epoch 6 compared to epoch 5 (698 ms), $F(1,12) = 152.21$, $p < 0.001$, $\eta_p^2 = 0.93$, along with a significant main effect of context (306 ms), $F(1,12) = 12.35$, $p < 0.01$, $\eta_p^2 = 0.51$. The epoch × context interaction was significant, $F(1, 12) = 4.84$, $p < 0.05$, $\eta_p^2 = 0.29$. One-way repeated-measures ANOVAs revealed no significant cueing effect in epoch 5 (147 ms), $F(1,12) = 1.29$, $p = 0.28$, $\eta_p^2 = 0.10$, but a significant effect in epoch 6 (466 ms; see previous paragraph). Similarly, the mean number of fixations was significantly reduced (by 1.88 fixations) in epoch 6 compared to epoch 5, $F(1,12) = 52.36$, $p < 0.001$, $\eta_p^2 = 0.81$, as well as (by 1.22 fixations) for “old” compared to “new” contexts , $F(1,12) = 10.79$, $p < 0.01$, $\eta_p^2 = 0.47$. The epoch × context interaction was also significant for fixation number, $F(1,12) = 7.28$, $p < 0.05$, $\eta_p^2 = 0.38$, reflecting absence of (reliable) contextual cueing in epoch 5, $F(1,12) = 0.67$, $p = 0.43$, $\eta_p^2 = 0.05$, but a significant effect in epoch 6 (2.01 fixations, see previous paragraph). The scanpath deviation ANOVA revealed a significant effect of context, $F(1,12) = 13.80$, $p < 0.05$, $\eta_p^2 = 0.54$, but not of epoch, $F(1,12) = 0.04$, $p = 0.84$, $\eta_p^2 = 0.004$, and the epoch × context interaction was significant, $F(1,12) = 11.44$, $p < 0.01$, $\eta_p^2 =0.49$. The interaction effect was mainly attributable to the absence of a (reliable) context effect in epoch 5, $F(1,12) = 0.14$, $p = 0.72$, $\eta_p^2 = 0.01$, but a significant effect in epoch 6 (see previous paragraph). Importantly, the scanpath deviation for old configurations was significantly reduced, by 3.27°, in epoch 6 relative to epoch 5, $F(1,12) = 4.98$, $p < 0.05$, $\eta_p^2 = 0.29$, which compares with an increase, by 2.92°, from epoch 5 to epoch 6 for new configurations, $F(1,12) = 8.37$, $p < 0.05$, $\eta_p^2 = 0.41$. The reduction of the number of fixations and the shortening of the scanpaths in epoch 6 (free-view display) relative to epoch 5 (tunnel view display) for "old" configurations suggests that viewing the search display under 8° tunnel vision conditions limits (optimal) scanpath planning during search performance. The significant lengthening of the scanpaths in epoch 6 relative to epoch 5 for "new" configurations is likely attributable to the increased number of available scanning choices with free-view compared to tunnel view displays. As this would equally have been the case with "old" displays, the more remarkable is the fact that the scanpaths were actually shortened (rather than lengthened) with "old" configurations. Finally, the mean fixation duration was significantly shorter in epoch 6 than in epoch 5 (29.47 ms), $F(1,12) = 10.54$, $p < 0.01$, $\eta_p^2 = 0.47$, while not being significantly influenced by context: main effect, $F(1,12) = 2.89$, $p = 0.12$, $\eta_p^2 = 0.19$; epoch × context interaction, $F(1,12) = 3.84$, $p = 0.07$, $\eta_p^2 = 0.24$. Overall, this pattern of results replicates the contrasts between the training session (with gaze-contingent viewing) and the transfer session (free viewing) for the two epochs that were most comparable in terms of time on the task and had, thus, the greatest potential for dissociating contextual learning and contextual retrieval.

Recognition session. In the recognition test, the mean hit rate (i.e., correctly identified old configurations) was 59.6%, which was only marginally higher than the false alarm rate (44.2%), $F(1, 12) = 4.20$, $p = 0.06$, $\eta_p^2 = 0.26$; *JZS Bayes factor = 0.87* (indicating that this effect cannot be regarded as "robust"). This marginal effect was largely due to two (out of a total of 13) participants; when those two participants were excluded, the difference between the hit rates and false alarm rates effectively vanished, $F(1,10) = 1.44$, $p = 0.26$, $\eta_p^2 = 0.12$, *JZS Bayes factor* = *2.52*. Thus, the results revealed no clear evidence of explicit learning of the repeated configurations. In addition, subject-wise analysis also revealed that the recognition sensitivity (*d'*) was not correlated with the magnitude of contextual cueing from the transfer session, $r = -0.13$, $p = 0.67$. This was consistent with previous study (Geyer, Shi, et al., 2010), which reported no correlation between explicitly recognizing a specific old display and the contextual cueing generated by that given display.

2.3.3. Discussion

Taking the results of both, the transfer and training, sessions together, we obtained a striking finding: although no contextual cueing was evident in the training session (not even in epoch 5), in which the viewing area was limited to a visible tunnel of 8° at any time during the search, a cueing effect manifested in the transfer session in which the search displays were always fully visible. This pattern suggests that search-guiding context is learned, but cannot be expressed with a gaze-contingent, limited-view display in which the global display layout is not available. Restated, being able to see only a limited, gaze-contingent tunnel area (of 8°) containing a local configuration of just 2-3 items can support contextual learning. However, effective retrieval of the acquired contextual associations requires additional information from the periphery (outside the 8° area), which likely contributes to optimizing (on-line) saccadic path planning and thus context-based search guidance. Experiment 2 was designed to examine whether such additional peripheral information would enhance contextual retrieval, by extending the gaze-contingent display from 8° (Experiment 1) to 12° (Experiment 2).

2.4. Experiment 2

2.4.1. Methods

The method was essentially the same as in Experiment 1, except that the size of the tunnel area was increased to 12° (thus including parts of the peripheral visual field), with a central, fully visible area of 7.5° and an outer transition ring covering 7.5°-12°. On average, 4.69 out of the total 12 items were visible at any given time. 13 participants (8 females, mean age: 24.08 years) with normal or corrected-to-normal visual acuity took part in the second experiment, after obtaining their informed consent. None of the participants were aware of the purpose of the experiment.

2.4.2. Results

The overall results are shown in Figure 2.4. Similar data analyses as in Experiment 1 were applied to both the RT and eye movement measures.

The overall mean error and outlier rates were low (errors: 0.90%; outliers: 3.03%). A two-way repeated-measures ANOVA of the error rates with context (old/new) and epoch (1-6) as factors revealed the main effect of epoch to be significant, $F(5, 60) = 3.66$, $p < 0.01$, $\eta_p^2 = 0.23$, but not that of the context, $F(1, 12) = 0.19$, $p = 0.67$, $\eta_p^2 = 0.02$, and the epoch × context interaction, $F(5, 60) = 1.28$, $p = 0.29$, $\eta_p^2 = 0.10$. The error rate decreased as the experiment progressed, indicative of reliable training and general (procedural) learning effects.

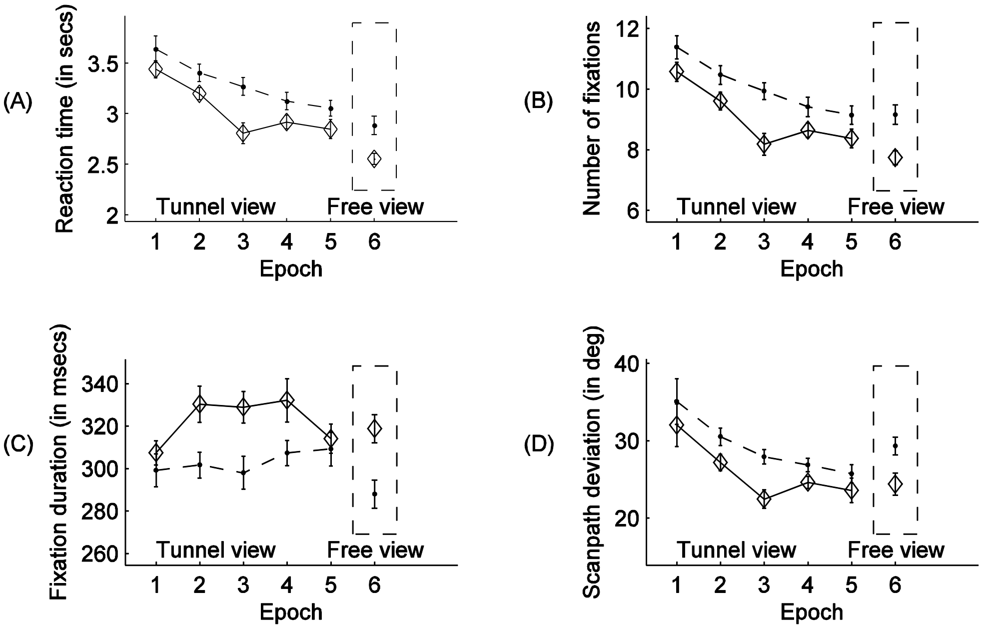

Figure 2.4 Results of Experiment 2. Mean reaction times (A), mean number of fixations (B), mean fixation duration (C), and scanpath deviation (D), all with associated standard errors, are shown as a function of experimental epoch and display context ("old", indicated by solid-diamond lines, vs. "new", indicated by dash-dot lines). Epochs 1-5 represent the training session with the tunnel view display, and Epoch 6 (in the dashed box) represent the transfer session with the free-view display.

Training session. Two-way repeated measures ANOVAs, with the factors context and epoch, were applied to the mean RTs, mean number of fixations, mean scanpath deviation, and mean fixation durations (the latter averaging 263 ms after refinement of the gaze samples). Mean RTs (Figure 2.4 A) changed across the training epochs, $F(4, 48) = 12.00$, $p < 0.001$, $\eta_p^2 = 0.5$, RTs were, on average, 589 ms faster in epoch 5 than in epoch 1, illustrating a typical effect of procedural learning. Importantly, in contrast to Experiment 1, RTs were significantly shorter for "old" compared to "new" configurations, $F(1,12) = 6.54$, $p < 0.05$, $\eta_p^2 = 0.35$, manifesting a typical contextual-cueing effect already in the training session (see comparison in Figure 2.5). The interaction between epoch and context was not significant, $F(4, 48) = 1.92$, $p = 0.12$, $\eta_p^2 = 0.14$, partly owing to variability (including non-reliable) of the cueing effects across epochs. The mean number of the fixations showed a similar pattern, as is illustrated in Figure 2.4 B. The mean fixation number decreased across the five training epochs, $F(4,48) = 14.34$, $p < 0.001$, $\eta_p^2 = 0.54$, with 2.22 fewer fixations in epoch 5 compared to epoch 1, and was reduced for “old” relative to “new” contexts (by 0.99 fixations), $F(1,12) = 8.08$, $p < 0.05$, $\eta_p^2 = 0.40$, without any interaction between the two factors, $F(4,48) = 2.03$, $p = 0.11$, $\eta_p^2 = 0.15$. These results indicate that both the general (procedural) learning and contextual-cueing effects, which are typically seen in classic contextual-cueing studies (without viewing restrictions), were also observable under conditions of gaze-contingent viewing when the tunnel size was extended from 8° to 12°. The mean scanpath deviation was also significantly influenced by epoch, $F(1.32, 15.87) = 5.53$, $p < 0.05$, $\eta_p^2 = 0.32$, being 8.88° shorter in epoch 5 than in epoch 1, and by context, $F(1,12) = 7.70$, $p < 0.05$, $\eta_p^2 = 0.40$, being 3.25° shorter for “old” compared to “new” displays; the epoch × context interaction was not significant, $F(1,12) = 1.46$, $p = 0.22$, $\eta_p^2 = 0.11$, indicative of a consistent advantage in saccadic scanning efficiency for “old” configurations in general. The mean fixations durations, depicted in Figure 2.4 C, exhibited no significant effects of epoch (main effect: $F(4, 48) = 1.16$, $p = 0.34$, $\eta_p^2 = 0.09$; interaction with context, $F(4,48) = 2.22$, $p = 0.08$, $\eta_p^2 =0.16$); but the main effect of context was significant, characterized by an increase, of some 20 ms, for “old” compared to “new” configurations, $F(1,12) = 5.69$, $p < 0.05$, $\eta_p^2 = 0.32$. This effect, which appears to run counter to the other measures, suggests that slightly extending the fixation duration may actually yield benefits in terms of improved saccade path planning.

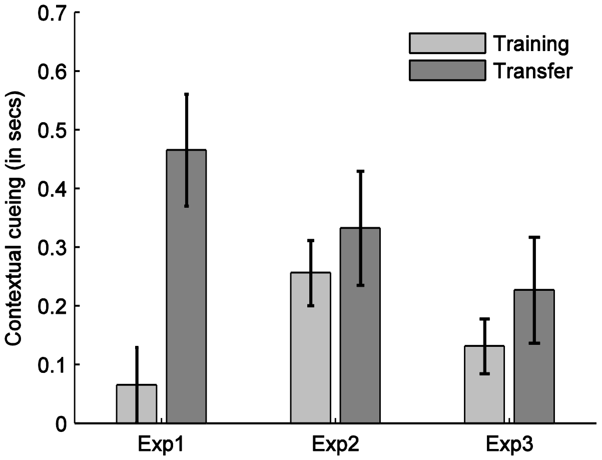

Figure 2.5. Comparison of contextual-cueing effects across experiments. The cueing effects in the training sessions were calculated by averaging across epochs 1–5.

Transfer session. As shown in Figure 2.4, significant contextual facilitation was evident for "old" compared to "new" configurations in the transfer session: RTs were shortened by, on average, 333 ms, $F(1,\ 12) = 11.67$, $p < 0.01$, $\eta_p^2 = 0.49$; fixation numbers were reduced by 1.42 fixations, $F(1, 12) = 14.63$, $p < 0.01$, $\eta_p^2 = 0.55$; scanpath deviation was decreased by 4.92°, $F(1,\ 12) = 14.59$, $p < 0.05$, $\eta_p^2 = 0.55$; and fixation durations were prolonged by 30.84 ms, $F(1,\ 12) = 10.04$, $p < 0.01$, $\eta_p^2 = 0.46$. These results suggest that "old" configurations enabled more efficient saccade path planning compared to "new" configurations.

To more closely examine whether a larger gaze-contingent tunnel area gives rise to a change of oculomotor and search behavior, epochs 5 and 6 (i.e., the last epoch of the training session and, respectively, the only epoch in the transfer session) were compared by means of two-way ANOVAs. The results revealed a significant drop of the mean RTs (231 ms) from epoch 5 to epoch 6, $F(1,\ 12) = 10.08$, $p < 0.01$, $\eta_p^2 = 0.46$, and a significant main effect of context (270 ms), $F(1,\ 12) = 10.66$, $p < 0.01$, $\eta_p^2 = 0.47$; but the interaction was non-significant, $F(1,\ 12) = 0.87$, $p = 0.37$, $\eta_p^2 = 0.07$. The ANOVA of the mean number of fixations revealed a significant main effect of context, $F(1, 12) = 10.95$, $p < 0.05$, $\eta_p^2 = 0.48$, with, on average, 1.01 fewer fixations for “old” than for “new” configurations; but neither the main effect of epoch (despite a slight drop in the number of fixations from epoch 5 to 6, by 0.31 fixations), $F(1, 12) = 0.93$, $p = 0.35$, $\eta_p^2 = 0.07$, nor the epoch × context interaction, $F(1, 12) = 1.66$, $p = 0.22$, $\eta_p^2 = 0.12$, were significant. Similarly, mean scanpath deviation was significantly smaller for “old” compared to “new” configurations (3.53°), $F(1, 12) = 8.57$, $p < 0.05$, $\eta_p^2 = 0.42$, while the effects of epoch, $F(1, 12) = 3.59$, $p = 0.08$, $\eta_p^2 = 0.23$, and the epoch × context interaction, $F(1, 12) = 2.40$, $p = 0.15$, $\eta_p^2 = 0.17$, were non-significant. Finally, the mean fixation duration was marginally longer, by 17.78 ms, for “old” than for “new” contexts, $F(1, 12) = 4.52$, $p = 0.055$, $\eta_p^2 = 0.27$, while the main effects of epoch, $F(1, 12) = 1.21$, $p = 0.29$, $\eta_p^2 = 0.09$, and the epoch × context interaction, $F(1, 12) = 3.39$, $p = 0.09$, $\eta_p^2 = 0.22$, were non-significant. Taken together, these results suggest that the faster responses in epoch 6 are likely attributable to continued procedural learning, rather than changes of oculomotor scanning engendered by the removal of the gaze-contingent viewing restrictions.

Recognition session. As in Experiment 1, participants’ mean hit rate (59.62%) was numerically higher than their mean false-alarm rate (51.92%), but this trend was non-significant, $F(1, 12) = 2.36$, $p = 0.15$, $\eta_p^2 = 0.16$, *JZS Bayes factor* = 1.73. In other words, there was no reliable evidence of explicit learning of spatial context, with the Bayes factor indicating that the null hypothesis (of implicit learning) was 1.73 times more likely to be true than the alternative hypothesis (of explicit learning).

2.4.3. Discussion

In contrast to Experiment 1, a contextual-cueing effect manifested in both the training and transfer sessions of Experiment 2, characterized by faster RTs, fewer fixations, longer fixation durations, and a smaller scanpath deviations for “old-” compared to the “new-context” configurations. These results indicate that a viewing window sized 12°, which makes 4-5 items visible at any given time, provides sufficient spatial information for both contextual learning and retrieval. Our results thus are consistent with previous findings that maintaining 3-4 local items invariant is sufficient for engendering robust contextual cueing (Brady & Chun, 2007). Considering the findings of Experiments 1 and 2 together suggests that contextual learning can be based on local configurations of only 2-3 visible items; however, being able to retrieve learned contexts for guiding search behavior requires additional, more peripheral (global) information, which likely helps optimize saccadic path planning.

If it is true that global information is necessary for contextual retrieval, then a brief preview of the global display prior to search under gaze-contingent, limited-viewing conditions might help engender contextual cueing. This prediction was examined in Experiment 3.

2.5. Experiment 3

2.5.1. Method

The settings of Experiment 3 were the same as in Experiment 1 (i.e., the gaze-contingent viewing area subtended only 8° of visual angle), except that a brief preview (150 ms) of the global display configuration was presented prior to the gaze-contingent search display (Figure 2.2). The preview display contained crosses (white crosses composed of two intersecting lines, each 1° in length) which marked the locations of all (forthcoming) search display items. That is, while being uninformative as to the identity of the (subsequently presented) search stimuli, the preview display provided information about the global context (structure) of the search display.

13 participants (11 females, mean age: 22.46 years) with normal or corrected-to-normal visual acuity took part in the experiment, after they had given informed consent. All participants were naïve with respect to the purpose of the experiment.

2.5.2. Results

Similar data analyses as in Experiment 1 were applied to both the RT and eye movement measures. The mean error and outlier rates were low (overall error rates of 0.82% and outlier rates of 2.61%, respectively). The error rates were comparable across epochs, $F(5, 60)=1.63$, $p = 0.17$, $\eta_p^2 = 0.12$, and contexts, $F(1, 12)=0.01$, $p = 0.91$, $\eta_p^2 = 0.01$, without any interaction between the two factors, $F(5,60) = 1.58$, $p = 0.18$, $\eta_p^2 = 0.12$.

Training session. Two-way repeated-measures ANOVAs, with epoch and context as factors, revealed the mean RTs to decrease significantly across the experimental epochs, $F(2.54, 30.51) = 5.62$, $p < 0.01$, $\eta_p^2 = 0.32$, reaching a reduction of 458 ms in the epoch 5 versus epoch 1 (Figure 2.6). The main effect of context was marginally significant (132ms), $F(1, 12) = 4.47$, $p = 0.056$, $\eta_p^2 = 0.27$, *JZS Bayes factor* = 0.952, without an interaction between epoch and context, $F(3, 36) = 0.95$, $p = 0.43$, $\eta_p^2 = 0.07$. This pattern of context effects is important as it suggests that a brief preview of the global display layout can bring back the contextual-cueing effect (see Figure 2.5), although contextual learning was slower and weaker compared to that in Experiment 2.

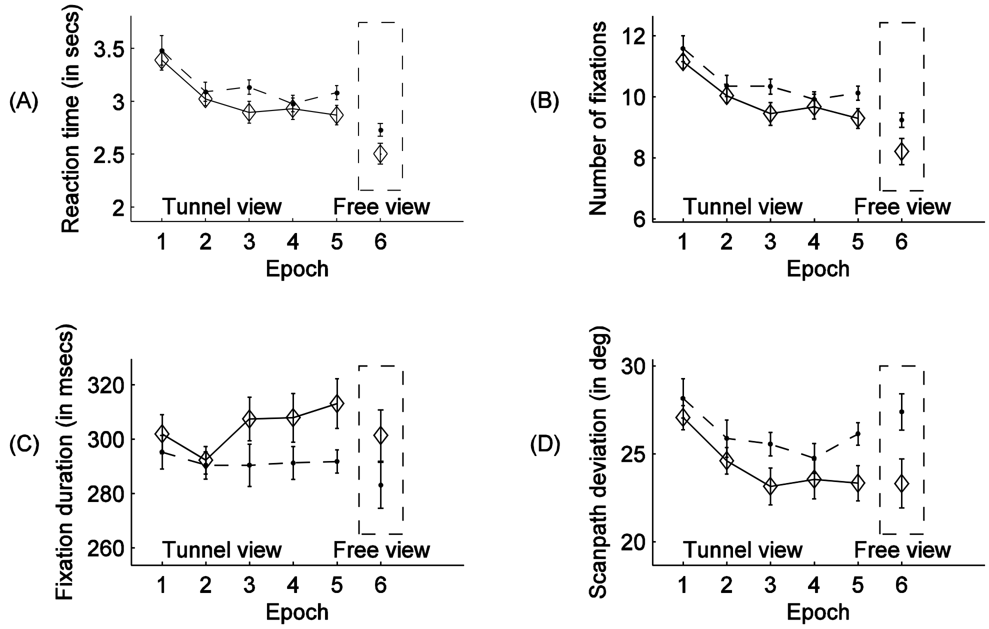

Figure 2.6. Results of Experiment 3. Mean reaction times (A), mean number of fixations (B), mean fixation duration (C), and scanpath deviation (D), all with associated standard errors, are shown as a function of experimental epoch and display context ("old", indicated by solid-diamond lines, vs. "new", indicated by dash-dot lines). Epochs 1-5 represent the training session with the tunnel view display, and Epoch 6 (in the dashed box) represent the transfer session with the free-view display.

In addition, the ANOVA of the mean number of fixations revealed both main effects to be significant: epoch, $F(2.86, 34.32) = 6.74$, $p < 0.001$, $\eta_p^2 = 0.36$, and context, $F(1,12) = 5.00$, $p < 0.05$, $\eta_p^2 = 0.29$. There were, on average, 1.66 fewer fixations in epoch 5 than in epoch 1, and 0.54 fewer fixations for "old" than for "new" configurations. The interaction between the epoch and context was non-significant, $F(4, 48) = 0.71$, $p = 0.59$, $\eta_p^2 = 0.06$. Interestingly, mean scanpath deviation was also significantly influenced by context, $F(1, 12) = 7.17$, $p < 0.05$, $\eta_p^2 = 0.37$, and epoch, $F(4, 48) = 3.14$, $p < 0.05$, $\eta_p^2 = 0.21$, without interaction between the two factors, $F(4,48) = 0.69$, $p = 0.61$, $\eta_p^2 = 0.05$: scanpaths were shorter for "old" than for "new" contexts (reduction by, on average, 1.75°), and in epoch 5 than in epoch 1 (reduction by, on average, 2.87°). Regarding the mean fixation durations, while there was no main effect of epoch, $F(4, 48) = 0.57$, $p = 0.68$, $\eta_p^2 = 0.05$, the context effect turned out significant: fixations were somewhat (12.71 ms) longer for “old” than for “new” contexts, $F(1, 12) = 7.08$, $p < 0.05$, $\eta_p^2 = 0.37$; the interaction between epoch and context as not significant $F(4, 48) = 0.79$, $p = 0.54$, $\eta_p^2 = 0.06$. Taken together, non-significant interactions in the oculomotor results suggest that contextual cueing may become effective early during the training, albeit being somewhat unstable.

Transfer session. Similar to Experiments 1 and 2, one-way repeated-measures ANOVAs analyses for mean RT, mean number of fixations, mean scanpath deviation, and mean fixation duration in epoch 6 (transfer epoch) revealed significant contextual effects for all measures examined: compared to the “new” configurations, the “old” contexts yielded significantly faster RTs (227 ms), $F(1, 12) = 6.33$, $p < 0.05$, $\eta_p^2 = 0.35$, while requiring fewer fixations (1.03), $F(1, 12) = 8.00$, $p < 0.05$, $\eta_p^2 = 0.4$, of slightly longer durations (18.30 ms), $F(1, 12) = 5.5$, $p < 0.05$, $\eta_p^2 = 0.31$, reaching the target via a shorter scanpath (4.07°), $F(1, 12) = 10.27$, $p < 0.01$, $\eta_p^2 = 0.46$.

Recognition session. The mean hit rate was 54.81% for repeated displays in the recognition session, which compares to a mean false-alarm rate of 46.15%; the difference between them was non-significant, $F(1, 12) = 3.02$, $p = 0.11$, $\eta_p^2 = 0.20$, *JZS Bayes factor* = 1.34. The associated Bayes factor suggests that the null hypothesis (of implicit learning) was 1.34 times more likely to be true than the alternative hypothesis (of explicit learning).

2.6. General discussion

Three experiments were conducted that combined gaze-contingent tunnel viewing of the search displays with a, in all other respects, standard contextual-cueing paradigm, in order to examine for differential roles of eye-centered local and peripheral-global spatial information in contextual learning and the retrieval of (learned) contexts. To assess whether the para-/foveal local information would be sufficient for contextual learning, the visible area of the gaze-contingent display (i.e., the tunnel area) in the training sessions (of 5 epochs) varied between 8° (including a 3° outer transitional belt) in Experiments 1 and 3 and 12° (including a 4.5° outer transitional belt) in Experiment 2, while a fully visible display was presented in the transfer sessions (epoch 6). In addition, a brief preview display was presented in the training session of Experiment 3 to examine whether brief exposure to the global display structure would aid contextual retrieval. When the tunnel area was limited to 8°, with on average only 2-3 items visible during any given fixation, no contextual facilitation was observed in the training session. Interestingly, however, contextual cueing was evident immediately (in the very first block of the subsequent test session) when the display was made fully visible - indicating that the repeated context was actually learned during the training session (Figure 2.5). When the tunnel area was extended to 12°, with on average 4-5 items visible during any given fixation, a robust contextual-cueing effect was obtained in both the training and transfer sessions, pointing to the need for (the availability of) peripheral global information for the cueing effect to become manifest. Experiment 3 further confirmed that global context information: even if made available only briefly (150ms) in a preview display that did not convey any fine-grained information as to item identity (i.e., whether a given item in the subsequent search display was a “T” or an “L”), it plays an important role for contextual retrieval.

It has been suggested that our perceptually coherent representation of the external environment is constructed based on the remapping of successive samples of local information onto their correspondent locations in a global spatial map (De Graef, 2007; Deubel, Koch, & Bridgeman, 2010; Intraub, 2002; Jonikaitis, Deubel, & de'Sperati, 2009; Melcher & Morrone, 2007; Schwarzkopf & Rees, 2010). Our findings suggest that humans are able to construct and learn repeatedly encountered spatial contexts based on scarce para-/foveal-local information (e.g., when just 2-3 items can be seen in each fixation), highlighting the role of local invariant inter-element spatial relations in contextual learning. Lack of para-/foveal-local information has been shown to impede contextual learning. For example, contextual cueing never developed for repeated displays with simulated foveal scotoma or fovea degeneration (Geringswald et al., 2012; Geringswald et al., 2013a). Peripheral global information, by contrast, affords fast processing of coarse, global-scale spatial information (e.g.,Hochstein & Ahissar, 2002; Navon, 1977; Schyns & Aude, 1994), which is important for saliency-based guidance and saccadic planning (Itti & Koch, 2001; Müller et al., 1995a; Wolfe, 1994b). The scanpath deviation results obtained in the present study also support the role of peripheral information in contextual retrieval and search guidance. Availability of some peripheral information when scanning the display under “wide” (12°) tunnel view conditions or brief (150 ms) availability of the display layout prior to the search led to a significant reduction of the deviation of the actual from the shortest (i.e., optimal) scanpath. The reduction of the scanpath deviation for the wide, but not the narrow, tunnel view suggests that peripheral information aids retrieving the learned spatial inter-element relations from contextual memory.

It should be noted that the roles of para-/foveal-local and peripheral-global information that we investigated here is different from the local/global manipulation in previous studies (e.g., Brady & Chun, 2007; Song & Jiang, 2005), in which “local information” often refers to the near-target local configuration, and “global information” to background features (e.g., color) or surrounding objects (Brockmole et al., 2006; Brooks et al., 2010; Kunar et al., 2006; Rosenbaum & Jiang, 2013). Both local and global spatial configurations were continuously available during search performance in previous studies. Using conditions of full display presentation, Brady and Chun (2007) found that when the repeated distractors (e.g., 2 “L”s) were locally positioned near the target, participants were able to acquire the context in the learning phase, suggesting that near-target invariant inter-element relations are important for contextual learning. Song and Jiang (2005), on the other hand, found that partial repetition – 1 "T" and 2 "Ls" randomly selected from the total 12 search items (not limited to the local area near the target, i.e., there could also be "long-distance" invariant relations) – was not sufficient to engender a contextual-cueing effect in the learning phase. But such partial repetition was important for contextual retrieval. Once a fully visible repeated display was learned, keeping only a part of the repeated display unchanged (e.g., 2 "L"s and 1 "T" out of the 12 items) was sufficient for maintaining contextual facilitation. The differential effects of near-target invariant (Brady & Chun, 2007) and "long-distant" invariant relations (Song & Jiang, 2005) can be well explained by the roles of para-/foveal-local and peripheral-global invariances in contextual cueing that were established in the present study. As para-/foveal-local information is the basic building block for constructing the whole spatial map, local invariances are important for successful contextual learning. Faster processing of peripheral spatial information (e.g., Hochstein & Ahissar, 2002; Navon, 1977; Schyns & Aude, 1994), on the other hand, does not require full matching of the global configuration (Song & Jiang, 2005); rather, some peripheral invariant features, such as "long-distance" invariances, are necessary for efficient search guidance and oculomotor planning. Such peripheral invariances can aid contextual learning and retrieval even if, during the actual search process, the global information is available only from working memory (WM), as shown in the present Experiment 3. Unlike the constant availability of peripheral information with free-view displays, though, just having a preview of the whole display (as in Experiment 3) would require additional processing to maintain the spatial configuration (in WM), which may be the reason for the relatively weak contextual-cueing effect that we obtained in Experiment 3. This, however, would require further investigation. Also, further work is required to establish whether the brief exposure of the global context strengthens the encoding of associations and/or facilitates the retrieval process. A recent study of contextual cueing in "pop-out" visual search suggests that the latter may be the case (Geyer, Zehetleitner, & Müller, 2010), while of course not ruling out the former. Geyer, Zehetleitner, et al. (2010) found that even supposedly automatic detection of salient pop-out stimuli may be facilitated by learning (i.e., long-term memory) contextual associations given that the global display configuration is exposed prior to the presentation of the search display proper. Geyer et al. argued that the preview permits contextual memory to be activated, to top-down influence the speed with which odd-one-out target is singled out by bottom-up feature contrast operations.

Considering oculomotor behavior, we observed that both procedural and contextual learning improve saccadic efficiency, in terms of a reduction of the number of fixations. Similar results have been reported in previous studies (Peterson & Kramer, 2001; Tseng & Li, 2004; Zhao et al., 2012), in which contextual facilitation was associated with a decrease in the number of fixations and saccades. While there is thus convergent evidence for RTs and the number of fixations to be reduced when searching through repeated (vs. newly encountered) configurations, it is still unclear, however, whether repeated contexts reduce or increase the mean fixation duration. For instance, Tseng and Li (2004) and Zhao et al. (2012) found no contextual facilitation of inter-saccadic fixation duration for repeated configurations. Van Asselen et al. (2011), by contrast, observed contextual cueing to be associated with shortened fixation durations, which they took to argue that repeated objects could be recognized faster than novel objects, thereby facilitating visual search. In our experiments, by contrast, we consistently observed prolonged fixation durations in conditions in which the contextual-cueing effect was observed. The strong coupling of extended fixation durations and contextual cueing suggests that the retrieved context alters oculomotor scanning so as to increase search efficiency. The finding of longer fixation durations is in line with several other visual search studies (Hooge & Erkelens, 1998, 1999; Zou, Müller, & Shi, 2012). For example, Zou et al. (2012) found that irrelevant tones that occurred during a hard visual search task led to an extension of the current fixation duration (during which the tone event occurred), increasing the likelihood of the subsequent saccade being directed to a display region that had not yet been inspected and thus potentially contained the target. Going beyond these visual search tasks, our finding confirms that extended fixation durations are coupled with reduced deviations of the actual scanpath (from the optimal, shortest path) for "old" configurations in contextual-cueing paradigms, reflecting improved planning of the saccade path (i.e., "contextual guidance") to the target location.

The gaze-contingent tunnel view conditions we realized here resemble the predicament of clinical patients with tunnel vision, whose sight is lost progressively from the visual periphery to the central fovea. One implication of our findings for such patients would be that they might be impaired with regard to contextual retrieval during the exploration of repeated scenes, but not necessarily with regard to contextual learning. In contrast to patients with tunnel vision, patients with age-related macular degeneration (AMD), who suffer from impaired foveal vision, have been shown to have difficulties in contextual learning (Geringswald et al., 2013a). A question that remains to be investigated in the future is whether the two patient groups do show differential deficits in contextual learning and retrieval. Such a comparison study would also be interesting with regard to patients with central versus peripheral vision loss developing differential search strategies to compensate for their deficits - as would be suggested by a recent study (Kwon, Nandy, & Tjan, 2013) which revealed the oculomotor system to be capable of adjusting saccadic behavior to compensate for a loss of foveal vision.

In conclusion, we demonstrated that repeated contexts can be learned based on limited local context (2-3 items) under gaze-contingent viewing of the search display, but cannot be effectively retrieved to aid search guidance. However, once (some) peripheral global information was provided or the whole display configuration was previewed, the contextual-cueing effect immediately manifested - suggesting that global information is necessary for contextual retrieval. The learned retrievable context was strongly coupled with changes of the oculomotor scanning patterns, indicative of better saccadic planning for learned and retrievable contexts.

Chapter 3. Transfer of contextual cueing in full-icon display remapping

3.1. Abstract

Invariant spatial context can expedite visual search, an effect that is known as contextual cueing (e.g., Chun & Jiang, 1998). However, disrupting learned display configurations abolishes the effect. In current touch-based mobile devices, such as the iPad, icons are shuffled and remapped when the display mode is changed. However, such remapping also disrupts the spatial relationships between icons. This may hamper usability. In the present study, we examined the transfer of contextual cueing in four different methods of display remapping: “position-order invariant”, “global rotation”, “local invariant”, and “central invariant”. We used full-icon landscape mode for training and both landscape and portrait modes for testing, to check whether the cueing transfers to portrait mode. The results showed transfer of contextual cueing, but only with the local invariant and the central invariant remapping methods. We take the results to mean that predictability of target locations is a crucial factor for the transfer of contextual cueing and thus icon remapping design for mobile devices.

Keywords: Contextual cueing, visual search, mobile interface, icon remapping

3.2. Introduction

Invariant visual context provides an important spatial cue for the guidance of visual search and focal-attentional selection. Repeated exposure to the same arrangements of search displays facilitates reaction time (RT) performance, an effect that has been referred to as contextual cueing (Chun, 2000; Chun & Jiang, 1998; Chun & Nakayama, 2000). In their seminal paper, Chun and Jiang (1998) had their observers search for a target letter “T” embedded in a set of distractor letters “L”. Unbeknown to participants, half of the presented displays contained identical configurations of target and distractor items (i.e., old displays), while the other half contained novel configurations (i.e., new displays). The main result was that of faster RTs to old relative to new displays (i.e., contextual cueing) – an effect that developed after a short period of training. Interestingly, when observers were queried about repeated displays at the end of the search task in an “old-new” recognition test, their performance was only at chance level. From these findings, Chun and Jiang (1998) concluded that (i) contextual cueing guides focal attention more rapidly to the target location, (but see Kunar, Flusberg, Horowitz, & Wolfe, 2007), for evidence that contextual cueing might also aid post-perceptual processes), and (ii) the cueing effect derives from an implicit memory for the items’ spatial arrangement. Since then, the cueing effect has been elaborated in a number of further studies (Chun, 2000; Chun & Jiang, 1998; Chun & Nakayama, 2000; Conci et al., 2011; Conci & von Muhlenen, 2009; Conci & von Mühlenen, 2011; Geyer, Shi, et al., 2010; Jiang & Wagner, 2004; Kunar et al., 2006). Jiang and Wagner (2004; see also Brady & Chun, 2007, or Olson & Chun, 2002) showed that contextual cueing is supported by two distinct spatial memory systems for individual item locations (i.e., local learning) and, respectively, the entire configuration formed by the distractors (i.e., global learning). Further, Kunar et al. (2006) showed that non-spatial attributes, too, such as background colour, can facilitate RT performance. Contextual learning is also influenced by selective attention: only the arrangement of some items, in particular, those sharing the target colour, are learned over the course of an experiment (e.g., Geyer, Shi, et al., 2010; Jiang & Leung, 2005).

However, the degree to which contextual cueing can adapt to changes in learned displays remains subject to debate. For example, Jiang and Wagner (2004) reported that contextual cueing was still reliable even when learned displays were shifted along the horizontal display axis, the vertical display axis, or presented in a different size (compressed or expanded). Other studies (Brady & Chun, 2007; Olson & Chun, 2002) showed that contextual cueing "survived" changes of approximately 50% up to 75% of the display items, that is, cueing was reliable even when only one half or one quadrant of the display was repeated across trials. On the other hand, Olson and Chun (2002) reported that the cueing effect was abolished when "new" distractors were presented in-between the target and the "old" distractors, with the target being presented, for example, in the left half and the "old" distractors in the right half of the display. Several other studies confirmed that contextual cueing diminished when the target was re-positioned in repeated displays and thus became unpredictable (Chun & Jiang, 1998; Manginelli & Pollmann, 2009; Olson & Chun, 2002). In contrast, the contextual cueing effect remained effective with predictable target location changes (Conci & Müller, 2012; Conci et al., 2011). Makovski and Jiang (2010) suggested that predictability based on invariant context is a key factor for contextual cueing, based on their finding that the cueing effect decreased as the target appeared further away from its "learned" location; in fact, there were even RT costs when the target swapped its location with a previous distractor. Similar findings have been reported in 3D-scene search (Chua & Chun, 2003), in which contextual cueing decreased with increasing angular difference between viewpoints in the training versus the test displays (the experiment was divided into a training and test phase, with the latter containing modified displays).

While most of the work on contextual cueing was conducted using consistent (i.e., spatially invariant) search displays with a fixed number of items (e.g., 1 target and 11 distractors presented at a total of 48 locations within an invisible 6 x 8 matrix), none of these studies has examined the influence of changes of the display orientation on the cueing effect. Although changing display mode (and accordingly remapping of the items) occurs rarely with "standard" (i.e., laboratory) displays, switching display mode is a normal routine in current touch-based mobile devices – such as the iPad. Interestingly, with these devices, there is only one type of item – or icon – remapping method available: the positions of icons in one display (e.g., landscape mode) are remapped to the other display (portrait mode) by keeping the positional order (left to right and up to down) constant across all icons (see Figures 3.1 A and 3.1 B). Although this remapping method preserves the positional order and 80% of the horizontal inter-icon relationships (in a 4 × 6 icon matrix, as shown in Figures 3.1 A and 3.1 B), it destroys almost all local icon relationships, in particular, when the display is arranged as a rectangle (as with almost all mobile devices). However, based on the contextual cueing studies reviewed above, it is possible that contextual cueing is reduced, if not entirely abolished, when display orientation changes from landscape to portrait mode and icons are remapped in the "standard" position-order manner. Given this, one intriguing question arises, namely: are there any other improved methods for icon remapping, such that the remapping could enhance users' performance in everyday situations of display mode changes? This question was addressed in the current study by using the contextual cueing effect as a tool to evaluate the effectiveness of various display remapping techniques, that is: preserved contextual cueing from one to the other display mode was taken as an indicator for the value of a given remapping method.

Besides the position-order remapping method, several other (simple) remapping methods are possible. For example, one of the most natural ways is to rotate the entire display by 90° in clockwise direction (individual icons are rotated 90° in counter- clockwise direction to keep their appearance constant; see Figure 3.1 C). Such a “global rotation” is similar to the rotation of an object in our physical world (e.g., imagine you rotate a key cabinet with many keys). Alternatively, and motivated by the above mentioned studies on contextual cueing (e.g., Brady & Chun, 2007), one could also try to preserve local associations within the entire configuration as completely as possible. There are two ways to maximize such local invariants. One is to subdivide displays into several local regions and preserve the placement of these local regions in the entire configuration after icon remapping (Figure 3.1 D). Another method is to keep the display centre constant in remapped displays (Figure 3.1 E).

In order to investigate how these various display remapping methods influence memory performance, we examined contextual cueing effects in four separate experiments. Each experiment examined one display remapping method. To simulate touch-based icon displays and observers’ active touch action, we used real desktop icons as search items and presented them on a touch monitor in the four experiments.

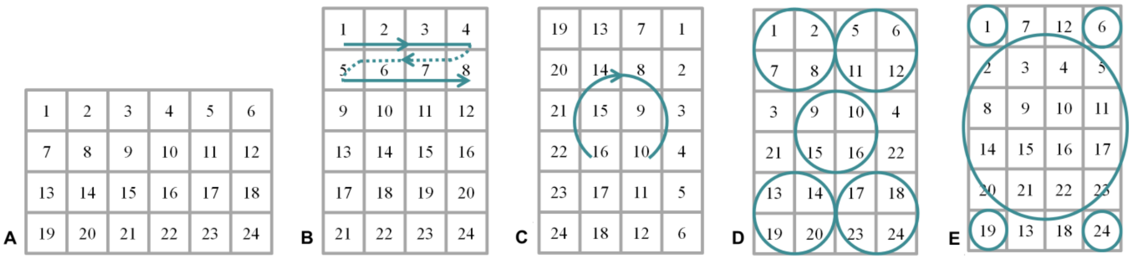

Figure 3.1 Schematic illustrations of display layouts and remapping methods. A) Display layout in landscape mode; each number denotes an individual icon. B) Portrait display layout obtained by the position-order invariant remapping method; the arrow indicates the icon remapping sequential order from the landscape to portrait mode. C) Portrait display layout obtained by the global rotation remapping method; the arrow indicates the rotation direction from the landscape to portrait mode. D) Portrait display layout obtained by the local invariant remapping method; circled regions remain the same between the landscape and portrait mode. E) Portrait display layout obtained by the central invariant remapping method; circled regions are invariant.

3.3. Experiment 4-7

3.3.1. Methods

Participants. A total of 40 observers took part in the experiments (10 in each experiment, mean ages: 27.9, 26.2, 25.5 and 27.3, number of females: 7, 6, 6 and 5 for Experiments 4-7 respectively). All had normal or corrected-to-normal visual acuity (including colour vision). They gave written consent prior to the experiment and were paid at a rate of 8 Euro/hour for taking part. Participants were naive as to the intention of the study.

Apparatus and stimuli. The experiments were conducted in a dimly lit cabin (ambient light: $4.36 cd/m^2$). Visual stimuli were presented on a 23-inch multi-touch LCD monitor (HP2310ti) with spatial resolution set to 1920×1080 pixels. In order to make touch-pointing comfortable for the participants, the screen panel was placed on the table tilted by 45°. The viewing distance was approximately 40 cm, with participants' head position fixed by a chin rest. 24 typical computer icons (randomly selected from 48 candidate icons with creative commons attribution copyrights for each observer) were presented within an invisible 6×4 horizontal grid (subtending 24°×16° of visual angle) or a 4×6 vertical grid (subtending 24°×16°). The target was the icon with a top overlay of a compound letter "T" (subtending, 1.6°×1.6°; luminance, $35.67 cd/m^2$; see Figure 3.2). Such a compound target letter was used for two reasons: first, to avoid interference between the target and some other (distractors) letters (e.g., the letter "S" in the Skype icon; Figure 3.2); second, to make the compound letter and the icon comparable in terms of their luminance level. The background of the search displays was set to grey ($16.56 cd/m^2$). To enhance the global spatial "Gestalt" (i.e., perception of the display as landscape or, respectively, portrait mode), we added one array of 6 up-right white triangles ($130.5 cd/m^2$) with a grey background ($19.62 cd/m^2$) below the landscape mode (Figure 3.2 A) or to the left side of the portrait mode (Figure 3.2 B). The triangle array was meant to serve as a global landmark in the experiments, indicating display mode changes. The experimental program was developed with and controlled by Mat lab (Mathworks Inc.), in addition to the Psychtoolbox (Brainard, 1997; Pelli, 1997). Response times were recorded via the touch screen. In order to determine the onset of a response, an additional input button (connected to a NI PXI system) was placed in-between the touch screen and the participants, which was used for initiating the task and pointing movement.

Design and procedure.A three-factorial within-subject design was used with display mode (landscape, portrait), context (old, new), and experimental epoch (1-9) as independent variables. From the 24 possible target locations, we randomly selected 12 target positions for old and the other 12 positions for new displays. In This way, the target appeared equally likely at any of the 24 possible locations. In order to have enough difference between old and new configurations and to control the similarity of icons identities, we selected 24 icons from 48 typical icon candidates and assigned to random locations. Each of the new target locations was paired with newly generated distractor icons for every new-display trial, whereas each of the old target locations was paired with randomly selected distractor icons at the beginning of each experiment and served as old landscape displays. These old landscape displays were also used to define the remapped old portrait displays. Remapping was one as follows:

(i) Experiment 4 (“position-order invariant”). The positional order (left-to-right and top to bottom) of the icons in the portrait mode was the same as that in the landscape mode (Figure 3.1 B). This method is used in most of the present mobile devices for the rearrangement of icons.

(ii) Experiment 5 (“global-rotation”). The landscape display was, as a whole, rotated by 90° clockwise into the portrait mode, while preserving the (upright) orientation of the individual icons. With this global rotation, the global and local relationships of the icons are rotated by 90° across display changes (Figure 3.1 C).

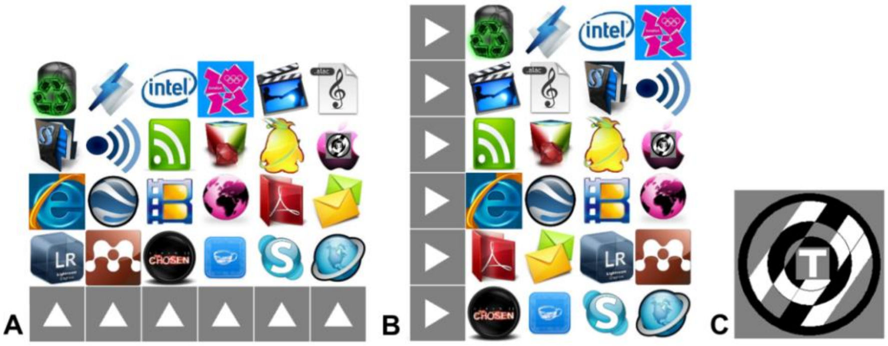

Figure 3.2 Example displays in the experiments. A) Example of a landscape display. In this example, the “Apple” icon (second row, right-most column) is the search target. B) Example of a portrait display. In this case, icons are remapped from the landscape mode by keeping the position order constant in the left-to-right and up-to-down manner (Experiment 4). C) The top overlay for the target icon (a compound letter “T”).

(iii) Experiment 6 (“local invariant”). To preserve the local (and global) spatial configuration as much as possible, in Experiment 6, the display was divided into four peripheral and one central region, each consisting of four icons (see circled regions in Figure 3.1 D). The positioning of these four “corners” and the central region were kept constant across display mode changes. Only four remaining items (i.e., icons 3, 4, 21, and 22 in Figure 3.1 D) changed their relative positions. Similar to the global rotation, with the local-invariant transformation, the local relationships between all icons are preserved across display changes.

(iv) Experiment 7 (“central-invariant”). Instead of dividing the display into multiple regions, in Experiment 7, we preserved the central display region as much as possible (i.e., preserving the central maximum square region). As shown in Figure 3.1 E, icons in the central 4x4 matrix were positioned at identical locations across display mode changes. In addition, the four outermost (‘corner’) icons were also unchanged. Only the remaining four icons (7, 12, 13, and 18 in Figure 3.1 E) changed their positions.

Each experiment comprised of three consecutive sessions: learning, test, and recognition. In the learning session, there were 5 epochs of 3 blocks, with each block consisting of 24 search trials. To keep the experiment as short as possible, learning session contained only 12 old-landscape displays to foster learning effect (each of the “old” display repeated twice per block). The test session had four epochs, with each epoch consisting of 24 trials (i.e., one block only). In half of these trials, an old display was presented, and new displays in the other half. New displays were randomly generated at the beginning of each trial. The order of display modes in the test epochs was fixed: landscape (L), portrait (P), portrait (P), and landscape (L). The first test epoch with the landscape mode (i.e., non-transformed) was intended to test for a “standard” contextual cueing effect. The last test epoch was intended for examining whether contextual cueing is still manifested by two intervening epochs containing different display modes. In order to avoid confounding by repetition effects, we randomly presented trials in such a way that the same old display was never repeated within 3 consecutive trials.

In the learning and test sessions, each trial started with a cross fixation presented in the centre of the display. Participants had to press the input button (also serving as the initial hand position) to trigger the presentation of the search display. Participants were instructed to detect the target and touch its location with their index finger as rapidly and as accurately as possible. A blank screen was presented after the localization response, or 4.5 seconds when no response was made. When participants made an erroneous response, an additional feedback display containing a stop warning sign was presented for 1.0 second. After 1.0 to 1.2 seconds of inter-trial interval, the next trial started.

In the recognition session, participants were asked if they had realized any display repetitions during the learning and test sessions and, if so, when they had first noticed the repeated displays (note that a similar protocol was used by Chun & Jiang, 1998). Following this, they had to judge a total of 24 displays, including 12 new displays (6 landscape and 6 portrait displays) and 12 old displays (6 landscape and 6 portrait displays), in an “old-new” recognition test. In this test, the chance rate for recognizing a repeated display was 50%.

Prior to the experiment, participants practiced the experimental task in one training block of 24 trials (data not recorded). The search displays used in the practice trials were not shown later in the experiment. Participants were allowed to take break in-between successive blocks of the experiment. The break between the learning and test sessions was similar to other between-block breaks.

3.3.2. Discussion

The findings of Experiment 1 suggest that context can be learned with a tunnel view of 2–3 display items per fixation, but the learned context cannot facilitate search (likely because contextual associations cannot be effectively retrieved) under such limited viewing conditions. The findings of Experiment 3 further indicate that a brief preview of the whole display configuration can help reinstate the contextual-cueing effect even if oculomotor scanning is subject to the same restrictions as in Experiment 1; that is, presumably, contextual associations may be retrieved in response to the preview display and thus guide scanning behavior. Taken together, the pattern of effects indicates that the global context, even if available only briefly prior to the search, plays an important role in the retrieval of contextual associations.

3.3.3. Results

Accuracy performance. Error rates were overall small (<1%) and were comparable across all experiments. For further RTs data analyses, we excluded trials with erroneous responses and RTs outside the range 200 to 3000ms. Such outliers were also low in general (<3%).

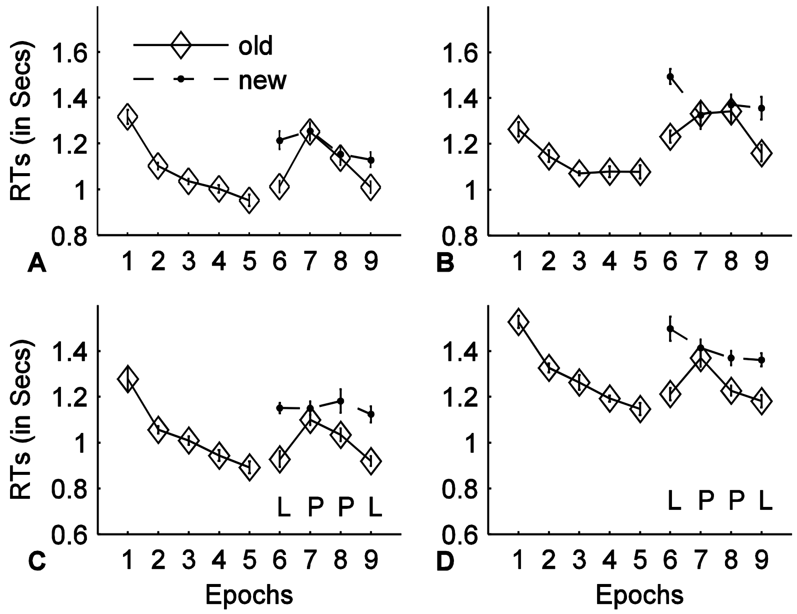

Figure 3.3 Mean correct RTs as a function of epoch for the learning (epochs 1-5) and test (epochs 6-9) sessions; for the latter, mean RTs are shown separately for old displays (denoted by diamonds and solid lines) and new displays (denoted by dots and dashed lines). A) Experiment 4, Position-order invariant remapping. B) Experiment 5, Global rotation remapping. C) Experiment 6, Local invariant remapping. D) Experiment 7, Central invariant remapping.

Perceptual learning.The mean RTs for the learning sessions are shown in Figure 3.3 (epochs 1-5). For each experiment, the mean RTs were examined by repeated-measures ANOVA with the single factor epoch. The main effect was significant for all four experiments (all “p”s<0.05); further Bonferroni tests revealed a significant perceptual learning effect, defined as the difference in RTs between epoch 5 (i.e., the end of the training session) and epoch 1 (i.e., the beginning of the training session) (see Table 3.1). In addition, to examine interference by the introduction of new (both landscape and portrait) displays in the test session, RTs for the old displays in the first epoch of the test session (epoch 6) were compared to RTs in the last epoch of the learning session. Although RTs were numerically longer in epoch 6 compared to epoch 5, the slowing was significant only for Experiment 5 (see Table

3.1 and Table 3.2). This suggests that introducing novel displays had only some moderate influence on the search task response.

Transfer of contextual cueing effect. The mean RTs, separately for old and new contexts, as function of epoch for the test phase are presented in Figure 3.3 (epochs 6-9). To examine the contextual cueing effect, mean RTs were subjected to a repeated-measures ANOVA with epoch (6-9) and context (old vs. new) as factors, separately for each experiment. The results are summarized in Table 3.2. RTs were significantly faster for old displays compared to new displays in all four experiments, indicating robust contextual cueing benefits. The main effect of epoch was also significant for Experiments 4, 6, and 7, indicating that some perceptual learning also occurred in the test session. Finally, the context × epoch interaction was significant for all experiments, reflecting differential cueing effects in the different epochs. Post-hoc tests revealed significant contextual cueing to be significant for all landscape displays (epochs 6 and 9). By contrast, for portrait displays (in epoch 8), significant contextual cueing was evident only in Experiments 6 and 7 (see Table 3.2). Note that each epoch in the test session contained only 24 trials, suggesting that the contextual cueing effect could be quickly transferred with the local invariant and central invariant remapping methods when the display mode was changed.

Table 3.1 Mean learning effect in the training sessions and interference by the addition of new displays in the test session, for each experiment

Experiment	Perceptual learning		Interference associated with the presentation of new displays	
	Facilitation (ms)	ANOVA	Cost (ms)	ANOVA
4	364	P < 0.01	60	p = 0.08
5	185	P < 0.05	154	p < 0.05
6	385	P < 0.01	36	p = 0.21
7	380	P < 0.01	98	p = 0.11

Table 3.2 Contextual cueing effects in the test session

Experiment	Contextual cueing effect (ms)					ANOVA test with F-value		
	Average (Epoch6~9)	Epoch6	Epoch7	Epoch8	Epoch9	Context(old/new)	Epoch (6-9)	Interaction
4	86	202***	5	18	120*	9.25*	9.23**	4.49*
5	120	262***	7	28	197**	12.48**	1.40	8.80***
6	156	223**	50	147*	205**	21.49***	5.16**	3.07*
7	163	286**	43	142*	190**	22.42***	5.63**	3.91*

*The reported significance values are as follows: *: p<0.05; **: p<0.01;***: p<0.001.*

To examine whether contextual cueing effects were comparable among the different experiments, a repeated-measure ANOVA was conducted on the cueing effect in the first test epoch 6 (with landscape mode), with the single factor experiment. The effect of experiment was non-significant, $F(3,27) = 0.55$, $p = 0.65$, suggesting that the contextual cueing effects were comparable among experiments. Thus, any differences in the subsequent test epochs are likely attributable to the particular method of display (icon) remapping.

Recognition test. Based on participants' post-experimental reports, we determined the percentages of participants who noticed display repetitions during the search task and who attempted to explicitly learn the displays; the times (in terms of the number of blocks performed) at which these participants first noticed the repetitions were also calculated. We then further calculated participants' mean hit and false alarm rates as well as their discrimination sensitivities (d') for landscape (L) and portrait (P) displays. The results are summarized in Table 3.3.

Table 3.3 Results of recognition test

Experi-ment	Noticed repetition	Explicit learning	When (blocks)	Hit rates	False alarms	d'_L	d'_P	$d'_L = d'_P$
4	90%	60%	6.89	76.3%	38.3%	2.17**	1.17*	p=0.17
5	80%	30%	7.25	72.9%	43.3%	1.33**	1.17*	p=0.68
6	90%	20%	5.85	65.1%	28.3%	1.56	2.02*	p=0.51
7	80%	60%	4.75	80%	28.3%	3.42**	1.52*	p<0.05

*The reported significance values are as follows: *: p<0.05; **: p<0.01.*

In all experiments, participants exhibited high proportions of recognized displays. The recognition sensitivities (d's) were significantly larger than zero for both landscape and portrait displays (p < 0.05), except for one marginally significant effect for the landscape display in Experiment 6 (p = 0.066), which was mainly due to one observer who showed an extreme negative d_H' score (-1.40).When excluding this participant, d_H' was also significant: p < 0.05. Taken together, the significantly positive d' scores suggest that after learning, participants recognized not only the old landscape displays, but also the remapped portrait displays in all four experiments. Moreover, there was no significant difference in recognition sensitivity between the landscape and portrait displays, at least for the first three experiments (see the last column in Table 3.3), indicating that remapping did not hamper explicit recognition. Although recognition accuracy was lower for portrait than for landscape displays in Experiment 7, the effect was mainly due to the very high recognition sensitivity in the landscape mode (see Table 3.3). Nevertheless, even in Experiment 7, the sensitivity for the portrait displays was still significantly greater that than zero, supporting the idea that the transformed "old" portrait displays can be recognized explicitly. The lack of differential recognition sensitivities between landscape and portrait displays in Experiments 4 and 5 is in contrast to the differential contextual cueing effects with landscape versus portrait displays.

This suggests that recognition and visual search may involve different memory processes, with the former recruiting more complex information matching processes that do not benefit the search processes.

3.4. Gernal Discussion

The present study examined the transfer of learned contextual cues in full-icon display remapping. The main goal was to investigate whether contextual cueing continues to facilitate icon localization (RT) performance following display mode changes. We compared four different types of icon remapping: position-order invariant, global-rotation, local-invariant, and central-invariant remapping. In all experiments, robust learning effects were found in the training session for the landscape displays. RTs were faster at the end relative to the beginning of the training session. This practice effect is likely attributable to general learning of the localization task (Schneider & Shiffrin, 1977).

In the test session, in which new displays were introduced (in addition to the old displays), we established a contextual cueing effects in all experiments – at least when the display mode was kept the same. This suggests that icon identities and spatial configurations among icons could serve as context cues to facilitate the localization task. Note, the facilitation effect might also be partially due to position-based learning, given that only “old” displays were used in the training session. However, the transfer effects found in the portrait displays (Experiments 6 and 7) cannot be fully explained by position-based learning, because the positions were changed in the portrait displays and positional repetitions were equated between the old and new displays. Interestingly, contextual cueing was evident for landscape displays even after the insertion of two epochs of portrait displays. This may be taken to indicate that the cueing effect is relative robust against interference within the same set of old configurations, consistent with previous studies (Chun & Jiang, 1998, 2003; Conci & Müller, 2012; Conci et al., 2011; Jiang et al., 2005; Song & Jiang, 2005; Zellin, Conci, von Mühlenen, & Müller, 2011). However, contextual cues acquired with landscape displays were transferred to portrait displays only under certain remapping conditions (those of Experiments 6 and 7), suggesting that contextual cueing is relatively inflexible and that transfer is confined to specific remapping situations.

The differential pattern of effects revealed among the four experiments raises the question as to the factors that modulate the transfer of learned displays. The position-order invariant method maintained icons in their same left to right and up to down manner. Although 80% of the horizontal relationships are preserved with this transformation, it destroys almost all vertical relationships. It also changes the absolute positions of the icons dramatically; for instance, position 5 is shifted from the left side in the landscape display (Figure 3.1 A) to the right side in the portrait display (Figure 3.1 B). As a result, the target location might become unpredictable in remapped displays, abolishing the contextual cueing effect (Conci et al., 2011; Manginelli & Pollmann, 2009). Note that, in the current terms, predictability refers to both the target’s absolute location on the screen as well as its placement within the entire configuration (given that we did not vary the target’s absolute and relative location independently).

When comparing the position-order invariant to the global-rotation method, the latter maintains all local icon neighbourhood relationships, but the overall “Gestalt” is rotated by 90° from the landscape to portrait mode. With this type of remapping, repeated displays failed to facilitate RT performance in portrait displays. Possibly, the contextual associations learned in the landscape displays were quite instance-specific and too weak for the global-rotation remapping. As shown in mental rotation studies (Böckler et al., 2011; Borst et al., 2011; Ionta & Blanke, 2009; Shepard & Metzler, 1971; Shomstein & Yantis, 2004). RTs increase linearly with increasing angular disparity when participants were asked to decide whether two presented objects are the same or not. Those paired objects were normally “rotated” objects or “mirrored” objects, and participants had to carry out mental rotation (rotating one object into the other) to solve the task. Applied to the current Experiment 5, although the global rotation maintains the local icons neighbourhood relationships, the mapping of a “new” portrait onto an “old” landscape display may likewise be a demanding (i.e., time-consuming) process, which diminishes any performance gains brought about by contextual cueing. In a previous study using 3D visual search, Chua and Chun (2003) also showed that contextual cueing decreased with increasing angular difference between viewpoints of training and test displays. Thus, demanding mental rotation might be the main reason why we failed to find any transfer of contextual cueing from the landscape to the portrait in Experiment 5. It should be noted, however, that in our setup, the experimental program presented the “rotated” portrait display automatically. That is, participants passively viewed the search displays, rather than carrying out the rotation actively. It would be interesting to examine the transfer of contextual cueing when participants rotate the displays themselves (i.e., actively).

In contrast to Experiments 4 and 5, we found significant transfer of contextual cueing in Experiments 6 and 7 - in which the portrait display was remapped from the landscape display using the local-invariant (Experiment 6) or central-invariant methods (Experiment 7). Both experiments disclosed numerical contextual cueing benefits already in the first epoch with portrait displays (50.8 and 44ms for Experiments 6 and 7, respectively), although these effects were not significant. No contextual cueing in the first portrait epoch is likely due to the orientation change of the whole display. Mapping “old” landscape to portrait displays may engage additional mental processes, diminishing the contextual cueing effect. In addition, inter-observer variability was large since both the “old” and “new” displays were presented only once in this epoch. Interestingly, transfer of contextual cueing was highly reliable for both remapping methods (147.4 and 142.0ms for Experiments 6 and 7, respectively) in the second epoch. The local-invariant remapping method keeps 5 out of 7 local regions unchanged and the global topological relationship of these 5 local regions also remains the same. This means that local regions appear at the very same positions (quadrants) in the entire configuration after the remapping. Likewise, the central-invariant remapping method maintains the absolute icon positions of the 4 outermost corners and the central region (83% in total). In both cases, after the remapping, the target position is much better predictable compared to both the position-order invariant and the global rotation methods. In contrast to previous investigations of contextual cueing, suggesting that only 3-4 repeated items (amongst some 8 novel items) can produce the effect (Song & Jiang, 2005), the results of the present Experiments 4 and 5 suggest that merely preserving some local invariant information does not guarantee transfer of contextual cueing. Instead, the 3-4 items would have to appear at the very same positions within the global configuration to observe contextual cueing (Experiments 6 & 7; see also Brady & Chun, 2007, for a related proposal, albeit using different approach).

The recognition tests showed that in all experiments, participants were well able to discern repeated from non-repeated displays. This contrasts with “standard” contextual cueing studies in which recognition accuracy was typically at chance level (Chun & Jiang, 1998). Explicit memory effects may be due to the heterogeneous and, importantly, realistic icons used as distractors in our experiments (see also Brockmole et al., 2006). Interestingly, in all the experiments of the present study, recognition accuracy was larger than chance for all landscape and, importantly, remapped portrait displays. In contrast, transfer of contextual cueing was observed only in Experiments 6 and 7. This argues that merely recognizing a repeated display as an old one does not necessarily mean that this also facilitates RT performance. Of interest in this regard, it has been reported that explicit learning of repeated displays engages neural processes that are distinct from those concerned with implicit configural learning (Geyer, Baumgartner, Müller, & Pollmann, 2012; Westerberg et al., 2011). Along these lines, we suggest that recognition and visual search are supported by different memory processes. Further, the dissociation between the transfer of contextual cueing (Experiments 6, 7) and explicit recognition (Experiments 4-7) suggests that the memory underlying explicit learning is more flexible than that underlying implicit configural learning.

In sum, our experiments suggest that when display orientation switches and icons are rearranged, the traditional position-order remapping method used in current mobile touch devices is suboptimal in aiding search performance. Comparing and contrasting three alternative methods of icons remapping, we found that when using local-invariant or central-invariant remapping, contextual cueing continues to enhance (target) icon localization performance. While the global-rotation method may be intuitive for users, it might introduce additional mental-rotation processes that are detrimental to localization performance. Our findings thus have implications for alternative interface design guidelines for icon rearrangement in mobile devices. Open questions awaiting further research concern how to optimize local invariance regions and what the effects of active manual rotation are.

Chapter 4. Foreground-background segmentation influences contextual cueing

4.1. Abstract

In previous two chapters, we found that maintaining the local information is critical for contextual learning. When the local information is scarce (Chapter 2), or reshuffled (Chapter 3), the contextual effect can hardly be archived. In Chapter 2, we also found that the availability of the global information is important for contextual retrieval. In this chapter, we will focus on the following question: whether the association of background context and foreground search array can influence contextual learning and retrieval. It has been shown that segmentation process influences the selective attention (Wolfe, 2003a), also selective attention could influence contextual learning (Jiang & Leung, 2005). However, it is still not clear whether foreground-background segmentation affects contextual cueing. In this chapter, we conducted three experiments to examine the dependency of item-independent context in contextual cueing on the foreground/ background segmentation. In the experiments, together with a classical contextual search display (a “T” and “L”s, we presented a task-irrelevant geometric cuboid frame, which was not changed during the training session, but was either rotated 90^o or removed in the subsequent transfer sessions. The cuboid was manipulated as the foreground information in Experiment 8, and as the background information in Experiments 9 and 10. The results showed that the contextual cueing effect was manifested in the training session across all experiments. However, the cueing effect diminished when the cuboid was changed or removed in Experiment 8, but not in Experiments 9 and 10. These findings suggest that segmentation process is prior to the contextual learning process. When the cuboid and search array were both selected in the learning phase, changes of cuboid can destroy the learned contexts. In contrast, when the cuboid was automatically segmented out from the search array, it has little influence on the contextual learning and retrieval.

Keywords: image segmentation, selective attention, connectedness

4.2. Introduction

In everyday life we are bombarded with huge amount of sensory inputs. To save mental efforts and cognitive resources, we benefit from sophisticated selective attention mechanisms that help us select the information that is important for performing a given task, while ignoring the irrelevant information (Treisman & Gelade, 1980; Wolfe, 2003a). This selective process builds up on prior knowledge or some heuristic cues that influences the deployment of attention. One classic example is image segmentation (Driver, Davis, Russell, Turatto, & Freeman, 2001a; Martínez et al., 2006).

The human visual system can effortless segment the observed visual input into candidate objects, that is grouping together those retinal inputs that are likely to correspond to the same object and separate from those that are parts of the other objects (e.g., Driver et al., 2001a). A number of studies have observed evidences that image segmentation occurred quite early in visual perception, yielding the candidate perceptual units (e.g., foreground task-relevant objects or information) for further attentional processing and other visual information (e.g., background task-irrelevant information) for inattentional processing (e.g., Baylis & Driver, 1992; Baylis & Driver, 1993; Driver et al., 2001a; Kim, 2013; Mazza, Turatto, & Umilta, 2005). For example, Mazza et al. (2005) have investigated the foreground-background segmentation and attention effect administering a change blindness task, in which participants had to judge whether any alternation (e.g., color variation) occurred between two consecutive displays (consist of foreground / background rectangles) that were intervened by a blank interval. The results revealed significantly higher detection sensitivity for the foreground change compared to the background change trials. Because of the foreground-background segmentation, participant’s attention was, by default, biased toward the foreground elements, therefore boosting better detection performance.

While on one hand the aforementioned studies observed the influence exerted by segmentation on deployment of attention, on the other hand, selective attention has been observed to change perception and cognitive behavior (e.g., Jiang & Chun, 2001; Jiang & Leung, 2005; Treisman & Gelade, 1980; Wolfe, 1994b; Wolfe, 2003a). For instance, it is understood that attended visual information is usually processed faster and more accurate compared to the information that is not attended (Daniel & Deubel, 2008; Kim, 2013; Mazza et al., 2005; O'Regan, Deubel, J., & Rensink, 2000). The behavior of selective attention has been meticulously observed in visual search tasks. For example the contextual cueing paradigm has proved a useful tool for exploring the facets of attentional selection (e.g., Chun, 2000; Chun & Jiang, 1998; Geyer, Shi, et al., 2010; Geyer, Zehetleitner, et al., 2010; Jiang & Swallow, 2013; Jiang, Swallow, et al., 2014; Jiang, Won, et al., 2014). In a standard contextual cueing paradigm, participants search for a “T”-shape target among a number of “L”-shape distractors. Unbeknown to the participants, half of the displays are repeated “old” displays (target/distractor’s locations were maintained across block) while the other half was “new” displays (items’ locations varied from block to block). These studies found that participants learned the repeated spatial configuration as indexed by faster search time for the “old” display compared to “new” display. Moreover the configural learning was rendered implicit, since participants were unable to tell apart “old” from “new” displays in a memory task. Altogether this robust observation is known as contextual cueing effect. The learning and expression of contextual cueing facilitation can be modified by different attentional deployment (Endo & Takeda, 2004; Jiang & Chun, 2001; Jiang & Leung, 2005). For instance, Jiang and Leung (2005) examined contextual cueing effect by asking participants to search for a white target among white and black distractors. After participants learned the constant spatial context (with both white and black distractors in constant locations), they were tested in a following transfer session with displays that maintained either the white (black varied) or the black (white varied) distractor sets. The results revealed that contextual cueing was enhanced with faster search speed when the locations of the target and the attended set (i.e., white distractors) but not the locations of the target and the ignored set (i.e., black distractors) were repeated, suggesting the selective attention modulates the expression of the contextual cueing.

Because image segmentation is a decisive factor upon the deployment of attention, we set-out to answer the question as to whether image segmentation interacts with the behavior of selective attention that can be expressed in the contextual cueing paradigm. To date, a number of studies have investigated item-dependent grouping and segmentation effect in contextual cueing (Conci et al., 2013; Conci & von Muhlenen, 2009; Conci & von Mühlenen, 2011), which observed that contextual cueing was significantly reduced for the grouped display with respect to the standard display. In their experiments, however, the visual search items were grouped and segmented by item's features (e.g., color, size and etc.); Although these studies have investigated the visual segmentation based on grouping factors, these finding are limited by the manipulation of item-dependent features, therefore leaving answers the question as to whether item-independent features, has an effect on the deployment of attention in the contextual cueing task. Furthermore, Conci and colleagues only focused on the gropuping effect of foreground context (i.e. visual search items), the different contributions of foreground / background context in contextual guided visual search is still unclear.

In order to investigate whether the item-independent foreground-background segmentation influences contextual cueing effect, the current study administrated three experiments presenting both visual search items ("T" and "L"s) and an additional pseudo 3D shape on the visual search stimuli. We aimed to answer whether item-independent feature information can be encoded together with the spatial configural context (i.e., target "T" and distractor "L"s), and whether the expression of contextual cueing depends on the characteristics of the item-independent features (i.e., foreground or background information) on the visual search display. We manipulated the features of the item-dependent information by constraining the positions and the depth of configural context and a pseudo cuboid shape information (i.e., a drawing cuboid shape which is presented on 2D plane, see Figure 4.1 as an example) in three experiments. The spatial shape was set as foreground in Experiment 8 and as background in Experiment 9 and 10. We first trained participants with visual search displays that include both constant configural information and the pseudo cuboid during the training session, and then either varied the cuboid or omitted it in the following transfer session. Therefore, based on the literature covered herein, we hypothesize that, should the pseudo cuboid (item-independent feature) be segmented in relation to the spatial configuration in terms of foreground / background associations during the training session, this should alter the way in which attentional selection is deployed in the task, thus the well-established contextual cueing effect should decrease or be abolished when the cuboid perspective changes or when it is absent in the transfer session. Furthermore we propose that contextual cueing transfer should be affected when the foreground but not the background item-independent cuboid changes in the transfer session.

4.3. Experiment 8

Experiment 8 investigated whether a foreground item-independent spatial shape (i.e., a pseudo cuboid) helped to group configural context and is encoded in contextual learning and expression. Crucially, we examined whether the already learned contextual cueing effect can be transferred, when the foreground shape varied or removed. In order to manipulate the item-independent cuboid as a foreground shape, we presented the visual search items only on the edges of the cuboid. In other words, visual search items were connected by the 9 lines that formed into the edges of cuboid (see Figure 4.1 as example). As proposed by previous studies

(Han, Humphreys, & Chen, 1999; Palmer, 1992; Palmer & Rock, 1994), a uniform connectedness between items (i.e., connect the “T” and “L”s in the current experiment) is a strong factor in perceptual organization and grouping which occurs at a very early stage. Therefore, the item-independent cuboid can be segregated together with the visual search items as foreground task-relevant information during visual perception.

4.3.1. Materials and Methods

Participants. Participants were given written informed consent in accordance with the declaration of Helsinki (2008), and were paid for their participation of 8 Euro per hour. Due to the reason that we are mainly interested in the transfer effect of contextual cueing when the spatial shape is varied, we only conducted the transfer session for those participants showing a positive contextual-cueing effect in the training session. A similar approach has been used in previous studies (Conci & Müller, 2012; Conci et al., 2011; Zellin, von Mühlenen, et al., 2013). In the end, 10 participants (7 females, mean age 26.5 years old) finished both the training and transfer sessions.

Apparatus and Stimuli. The visual stimuli were presented on a 21-inch LACIE CRT monitor with a refresh rate of 100 Hz in a dark cabin (2.95 ± 0.95 cd/m^2). The viewing distance was maintained to 57 cm by the support of a chin rest. The visual display was presented on a grey background (6.33 ± 0.91 cd/m^2), which consisted of two types of context: configural context and spatial shape context (see Figure 4.1). The configural context comprised of 12 visual search items (each item subtends as 0.8° x 0.8° visual angle, 24.24 ± 3.39 cd/m^2), which included one “T” shape target and eleven “L” shape distractors. Similar to previous studies (Jiang & Chun, 2001; Olson & Chun, 2002; Zang, Jia, Müller, & Shi, 2013), the “L” distractor had a small offset (0.12°) at the line junctions making the “L”s more similar to the target stimuli. The spatial shape context was a pseudo cuboid (i.e., a cuboid that is drawing on a 2D plane, see Figure 4.1 for an example), which composed of 9 white lines (24.24 ± 3.39 cd/m^2, with 12°x12° of visual angle). Two type of the cuboid was used in the present experiment: “up-tilted” pseudo cuboid and “down-tilted” pseudo cuboid. The square face of the “up-tilted” pseudo cuboid located in quadrant I (the right-up quadrant, see Figure 4.1, the cuboid used during the training session), while the “down-tilted” pseudo cuboid was formed by rotating the “up-tilted” pseudo cuboid 90° clockwise (see Figure 4.1, the pseudo cuboid used during the transfer session).

In each search display, “L” shapes could appear in one of the four orthogonal rotations (rotated as 0°, 90°, 180° or 270°), while “T” shapes were rotated 90° either clockwise or counter-clockwise, pointing to the right or to the left (requiring a “left” or, respectively, “right” response – see below). Both “T” and “Ls” were randomly located in a 36 possible spatial locations inside an invisible 11x11 grid square area, each location subtended as 1.2° x 1.2° of visual angle. The 36 possible item locations were selected on the edges but not vertex of the trained “up-tilted” cuboid (see the left panel in Figure 4.1A). In this way, the pseudo cuboid was tightly related to the configural items, therefore, serves as foreground task-relevant information.

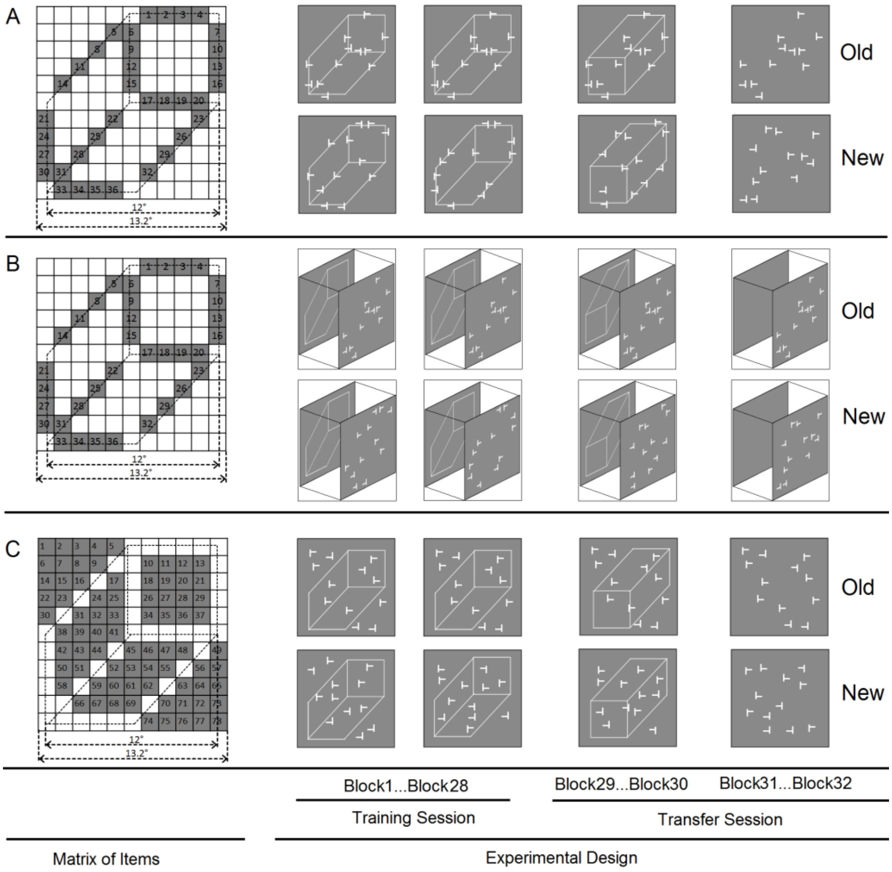

Figure 4.1 Stimulus configurations and schematic paradigm used in Experiments 8, 9 and 10. Left panel: Possible positions (gray grids) for search items in three experiments respectively. The visual search items were presented on the edges on the pseudo cuboid in Experiment 8 and 9 (A, B), however spread over the whole display except on the edges of the cuboid in Experiment 10 (C). The grids, numbers, and the grey color were invisible during the actual experiments. The background cuboid was inside an 11×11 matrix area, which subtended as 13.2°×13.2°. Right panel are schematic illustration of three sessions used in the experiments: the training session (block 1-28), the first transfer session (block 29-30), and the second transfer session (block 31-32). For "old" item-based configurations, each target was paired with a particular consistent distractor sets, repeated once per block; while for the novel item-based configuration, the target was paired with newly generated distractor sets for each presentation. The global background, however, was the same for both repeat and novel displays, with the "up-tilted" cuboid in the training session, "down-tilted" during the first transfer session, and was removed during the second transfer session. Of note, the visual stimuli used in Experiment 9 (B) was similar as the stimuli used in Experiment 8 (A) except the pseudo cuboid was presented on different depth plane compared to the configural context. The display was plotted from a -45° of view angle in order to show the depth information well. In the real experiment, however, participants view the display from the front of the visual search items.

Procedure and design. Participants were asked to discriminate the orientation of a target letter "T" among 11 distractor "L"s on the display as fast and accurate as possible by pressing a key, either the left or the right arrow key on the keyboard, using their index fingers. Each trial started with the presentation of a central fixation cross for 800-1000ms, which immediately followed by the search display. The search display remained on the screen until a response was made or the presentation exceeded 10 seconds. The next trial automatically started after a random interval of 1.0-1.2 seconds. As shown in Figure 4.1A, the experiment consisted of a 28-block training session, two 2-block transfer sessions and a 1-block recognition session. Each block of 16 trials contains 8 "old" and 8 "new" displays, randomly intermixed with each other. For the "old" displays, the configural context (with one "T" shape target and eleven "L" shape distractors) was consistent in both identities and locations, repeated once per block. For the "new" displays, however, the configural context varied by coupling a particular target location with newly generated distractor sets for each presentation. Importantly, the orientation of the target was randomly chosen in each trial for both "old" and "new" displays, so that the target's orientation itself won't lead to any context facilitation between "old" and "new" context.

During the training session, an "up-tilted" foreground pseudo cuboid, with visual search items presented on its edges (but not vertexes), served as the task-relevant scene for both "old" and "new" displays. Since the cuboid was identical, it cannot cue the target's location better for the "old" compared to the "new" displays. Any differences in search behavior between the "old" and "new" displays should either be due to configural context ("old" vs. "new"), or due to the interaction of task-relevant cuboid and the search items. In a subsequent transfer session, the task-relevant cuboid was rotated 90° clockwise into a "down-tilted" cuboid for both "old" and "new" displays to examine whether the learned contextual cueing, if any, can be transferred to the displays regardless of the change of the foreground pseudo cuboid. The configural context was maintained constant across different sessions for the “old” displays. With this variation, most of the visual search items (more than 88%) were not on the edges of the “down-tilted” cuboid, thus clearly destroys the association between the task-irrelevant cuboid and the search array. In a second transfer session, however, the foreground pseudo cuboid was removed (“no-cuboid”), served as the baseline, to investigate if the spatial context can be learned and retrieved in depend of the landmark information.

Once the search task was completed, participants performed a block of recognition task, which includes original 8 repeated and another 8 newly generated displays, with both types of the display presented without pseudo cuboid in order to examine whether participants had learned the configural context explicitly or not. Each display was presented for 20 s, and participants were informed that around half of the display was repeated display and the other half was new display. The task was to decide whether they had seen the current given display in the former visual search task or not, by manual key press input (left arrow key for “yes” and right arrow key for “no”). Prior to the experiment, participants practiced the search task with “up-tilted” cuboid in one block of 16 trials. The configural context of search displays used in the practice trials were not shown later in the experiment. Participants were allowed to take a break in-between successive block of the experiment.

4.3.2. Results

Search task. To increase the power of statistical analysis, every 7 consecutive blocks in the training session were grouped into one epoch, forming epochs 1-4, and every transfer session (2 blocks) was grouped into one epoch, forming epoch 5 (i.e., transfer session I) and epoch 6 (i.e., transfer session II) respectively. Trials with erroneous responses or reaction times (RTs) that were outside the range 0.2s and 3 standard deviations were excluded for further analysis. Both the overall mean error rates and outliers were low (mean error rates:

1.04%; outliers: 2.29%). The error rates were comparable across all conditions: context, $F(1, 9) = 2.22$, $p = 0.17$, $\eta_p^2 = 0.20$, epoch, $F(2.03, 10.73) = 1.92$, $p = 0.18$, $\eta_p^2 = 0.18$, and interaction, $F\ (2.26, 9.53) = 1.61$, $p = 0.22$, $\eta_p^2 = 0.15$, showing no improvement in accuracy by training. The overall results of the 10 participants are presented in Figure 4.2.

In the training session, repeated-measures analysis of variance (ANOVA) for reaction time (RT) with context ("old" vs. "new") and epoch (1~4) as factors, revealed significant context effect, $F(1, 9) = 8.48$, $p < 0.05$, $\eta_p^2 = 0.49$, with 213 ms faster for the old displays compared to the new displays, as well as epoch effect, $F(1.21, 10.92) = 22.36$, $p < 0.01$, $\eta_p^2 = 0.71$, with 340 ms faster in epoch 4 compared to epoch 1, but no interaction effect, $F(3, 24) = 2.47$, $p = 0.08$, $\eta_p^2 = 0.22$.

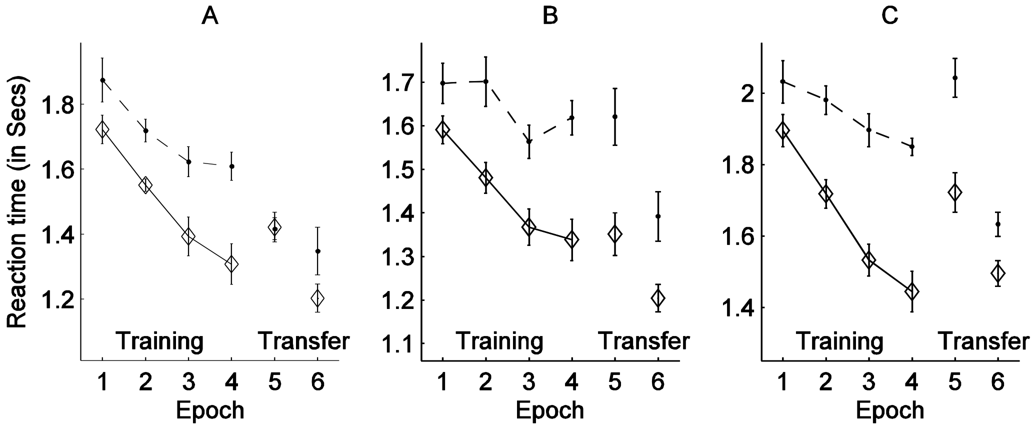

Figure 4.2 Results of the three Experiments. A) The results of Experiment 8 when the visual search items were presented on the edges of the cuboid. B) The results of Experiment 9 with 3D disparity vision; C) The results of Experiment 10 when the visual search items were presented on the empty space of the cuboid. Mean reaction times with associated 95% confidence interval, are shown as a function of experimental epoch and display context ("old", indicated by solid-diamond lines, vs. "new", indicated by dash-dot lines). Epochs 1-4 represent the training session, Epochs 5 and 6 represent the two transfer sessions respectively.

In the transfer sessions, participants tend to show faster response for old display compared to new display, however, this trend was neither significant in epoch 5 nor in epoch 6, manifested by one-way repeated measures ANOVA analysis: in epoch 5, F (1, 9) = 0.008, p = 0.93, η_p^2 = 0.001, with mean effect of -5ms; in epoch 6, F(1, 9) = 1.96, p = 0.20, η_p^2 = 0.18, with mean effect of 145ms. In addition, JZS Bayes factors (null / alternative) (Rouder et al., 2009) were calculated for epoch 5 and 6 to check whether our results favored the null hypothesis. According to Jeffries (1961), a value greater than 3 provides solid evidence for choosing the null hypothesis. The JZS Bayes factor was 4.29 for epoch 5 supporting the null hypothesis of the absence of contextual cueing, and was 1.85 for epoch 6 which provide "some" evidence for the absence of contextual cueing effect. The less solid evidence of null hypothesis and the trend of contextual facilitation in epoch 6 were mainly due to two out of ten participants who were significantly faster for the "old" configuration compared to the "new" configuration (> 700 ms). When excluded these two participants, there was no significant contextual facilitation any more, F(1, 7) = 0.07, p = 0.20, η_p^2 = 0.18, *JZS Bayes factor* = 3.79, mean effect of 7 ms, suggesting the well-established contextual cueing cannot be preserved for most of the participants when the foreground cuboid was removed. In summary, these results suggest that although the cuboid itself does not indicate target's location, the item-independent cuboid is integrated together with the configural context during contextual learning. Therefore, when the identity of the foreground cuboid was changed or removed, the context facilitation disappeared.

Recognition results. Those trials with response exceeded 20s were excluded for further analyses. In total only 0.42% of data was discarded. Participants' mean hit rate (i.e., correctly identified old configuration) was 60.54%, which was numerically higher than the false alarm rate (48.75%). However, this difference was not significant, F(1, 9) = 1.80, p = 0.21, η_p^2 = 0.17. The JZS Bayes factor (1.98) provided "some" but not strong evidence for the null hypothesis, suggesting participants cannot recognize the repeated configuration explicitly. Further subject-wise analysis revealed that the recognition sensitivity (d') was not correlated with the magnitude of contextual cueing in the last epoch of the training session, $r = 1.63$, $p = 0.65$, suggesting explicit memory (if any) is not a determination of contextual cueing effect.

4.3.3. Discussion

When the visual search array were presented on the edges of the pseudo cuboid, forming a foreground task-relevant shape during the training session, the learned contextual cueing can hardly be retrieved when the cuboid was rotated or removed during the transfer session. Of note, the missing of contextual cueing during transfer session may due to two possible reasons: on the one hand, the learned spatial context memory during the training session was constructed in part with reference to the foreground pseudo cuboid although the cuboid did not provide any informative spatial information about the target's location. In other words, the foreground cuboid was encoded during contextual learning. Given that the visual search items are linked by the edges of the foreground cuboid, which lead participant's spatial context representation to a cuboid- or partial cuboid like object. As a result, the learned contextual cueing is sensitive to the change of the spatial pseudo cuboid. In this case, if we can dissociate the tight relationship between visual search items and the cuboid in further experiments (i.e., a background cuboid), the variation of cuboid in the transfer session won't impede contextual cueing. On the other hand, the contextual learning may be independent of the pseudo cuboid that is the spatial context memory was constructed merely based on the configural context. However, the variation of the cuboid during the transfer session may block participant's ability to retrieve the learned spatial memory therefore impeded the learned contextual effect. If this is the reason that we observed no contextual facilitation during the transfer session of Experiment 8, the rotation or omission of the cuboid would always block participant’s ability to retrieve the well-established cueing no matter this cuboid is presented as foreground or background. In addition, what still equivocal with respect to the current experiment are: are all types of pseudo 3D shapes that are presented together with configural context on the visual search display were encoded during contextual-guided visual search? Do the contributions of the pseudo 3D shapes in contextual cueing depend on their properties (e.g., foreground vs. background information)? Due to the reason that attention is biased towards the foreground element (e.g., Mazza et al., 2005), and it is indispensable for successful contextual learning and expression (Jiang & Leung, 2005), a background task-irrelevant shape may play a different role in contextual cueing in comparison with foreground task-relevant shape information. In order to investigate these questions, we designed a second experiment which manipulated the pseudo cuboid as background task-irrelevant information on the display.

4.4. Experiment 9

Experiment 9 is designed to address the question whether the encoding of 3D shape in contextual learning is limited to the foreground task-relevant shape, or can encompass background task-irrelevant shape. The method was essentially the same as in Experiment 8, except 3D layout formulated by binocular disparity vision was adopted as the visual search stimuli. In order to generate the 3D layouts, Experiment 9 was conducted in a different lab cabin mounted with a 3D projector inside. Hence, some minor differences of the visual stimuli and apparatus between Experiment 8 and 9 exist, and they are described in the following part.

4.4.1. Materials and Methods

The configural context (includes one “T” and eleven “L”s) was still presented at the 36 possible item locations that are limited to edges but not the vertex of the pseudo cuboid (see Figure 4.1 B as an example). Nonetheless, the pseudo cuboid and the configural context were presented in different depth planes, providing configural context on the fixation plane (i.e., the frontal plane) and the cuboid on the back plane. As mentioned by (Nakayama & Silvermann, 1986) that human are able to perform a parallel search in one depth plane without interference from visual information in another depth plane, the background task-irrelevant cuboid in the current experiment may be sequestered from the configural context during contextual learning. Therefore, the variation of the pseudo cuboid may not decrease the well-established contextual facilitation.

Participants. Similar as in Experiment 1, only the participants who showed positive contextual-cueing effect in the training session were recruited in the transfer session. In the end, 10 participants (8 females, mean age 26.1 years old) finished the whole experiment, including both the visual search and recognition task.

Apparatus and Stimuli. The visual stimuli were presented by a 3D compatibility Optoma projector (HD131Xe) to a white canvas. The refresh rate of the projector was set as 120 Hz, and the display mode was set as 3D SBS (side by side). Note that the projector’s refresh rate was divided into left/right eye by Optoma ZF2100 3D glasses, and the final refresh rate for participant’s left or right eye was 60 Hz. The experiment was conducted in a dark cabin (13.2 ± 6.67cd/m^2) and the viewing distance was maintained to 77 cm by the support of a chin rest. The fixation plane (i.e., the frontal plane) was defined by a grey rectangle area (56.78 ± 1.23cd/m^2, with a size of 16°×16°, see Figure 4.1 B, right panel as examples) which enclosed either configural context (i.e., the visual search stimuli) or the fixation cross (i.e., the fixation display). This grey rectangle area was always available during the experiment. The configural context included a “T” and eleven “L”s (0.8° × 0.8° in visual angle, $97.62\pm 2.35 cd/m^2$) was presented on the center area (with a size of 13.2° × 13.2°) of the frontal fixation plane. A pseudo cuboid composed of nine lines ($97.62\pm 2.35 cd/m^2$) was presented on a back plane with a visual distance around 6 cm hinter the fixation plane when participants visual the display with 3D disparity glasses. Importantly, although the configural context and the pseudo cuboid were presented on different depth planes, the visual search items were visually on the edges of the pseudo cuboid. With these experimental settings, the pseudo cuboid was manipulated as background task-irrelevant cuboid which can be easily segregated from the configural context.

4.4.2. Results

Search task. Similar to the analysis used in Experiment 8, the trials were grouped into epochs 1-4 for the training session, and epochs 5 and 6 for the transfer session. Trials with erroneous responses or reaction times (RTs) that were outside the range of 0.2s and 3 standard deviations were excluded for further analysis. Both the error rates and the ratio of outliers were low (mean error rates: 1.58%; outliers: 2.96%). The error rates were comparable across all conditions: context, $F(1, 9) = 0.39$, $p = 0.55$, $\eta_p^2 = 0.04$, epoch, $F(5, 45) = 1.57$, $p = 0.19$, $\eta_p^2 = 0.15$, and interaction, $F(5, 45) = 0.48$, $p = 0.79$, $\eta_p^2 = 0.05$, showing no improvement in accuracy by training. The overall results are presented in Figure 4.2 B.

In the training session, repeated-measures analysis of variance (ANOVA) for reaction time (RT) with context (“old” vs. “new”) and epoch (1-4) as factors, revealed significant context effect, $F(1,9) =17.94$, $p < 0.01$, $\eta_p^2 = 0.67$, with 201ms faster for the old displays compared to the new displays, as well as epoch effect, $F(1.46, 13.13) = 6.00$, $p < 0.01$, $\eta_p^2 = 0.40$, with 166ms faster in epoch 4 compared to epoch 1, and interaction effect, $F(3, 27) = 3.13$, $p < 0.05$, $\eta_p^2 = 0.26$. The significant interaction effect was mainly due to marginally significant differences between old and new contexts during the first epoch, F(1, 9) = 4.77, p = 0.06.

In the transfer sessions, one-way repeated measures ANOVA showed significant contextual cueing effect in both epoch 5, F(1, 9) = 7.63, $p < 0.05$, $\eta_p^2 = 0.46$, with 269ms faster for old context compared to new context, and epoch 6, F(1, 9) = 7.45, $p < 0.05$, $\eta_p^2 = 0.45$, with 188ms faster for old context than new context. These results suggested that spatial context can be learned independent of the pseudo cuboid during the training session when it was presented on a different depth plane compared to the configural context. Although the configural items were still presented on the edges of the cuboid (same as experiment 8), the learned contextual cueing was maintained when the cuboid was changed or removed.

Recognition results. All the trials were responded within 20 s. The mean hit rate was 46.24%, which revealed no significant difference from the false alarm rate (52.5%), $F(1, 9) = 1.0$, $p = 0.34$, $\eta_p^2 = 0.1$, *JZS Bayes factor* = 2.74, suggesting participants can not explicitly recognize the old configuration after the executing of the visual search task under 3D condition. Further subject-wise analysis revealed that the recognition sensitivity (d') was not correlated with the magnitude of contextual cueing in the last epoch of the training session, $r = 1.64$, $p = 0.65$.

4.4.3. Discussion

In this experiment, the configural context and the item-independent cuboid were presented on different depth planes on the visual display. Although the search array was still presented on the edges of the pseudo cuboid, it can be easily segregated from the configural context. Interestingly, we observed the learned contextual cueing with “up-tilted” cuboid in the training session was maintained when this cuboid was rotated or removed in following transfer sessions. These results suggested that the spatial configural context can be learned independent of the background task-irrelevant cuboid, and the variation of the cuboid itself won’t block contextual retrieving. Due to the reason that the visual stimuli in Experiment 9 were very similar to the stimuli in Experiment 8, by keeping the similar visual item arrangement (i.e., all of the visual search items were presented on the edges of the pseudo cuboid), the foreground-background segmentation effect can be the crucial factor that leads to different contributions of the pseudo cuboid in contextual learning. In other words, the repeated spatial context can be learned independent of the background task-irrelevant cuboid but with reference to the foreground task-relevant cuboid. Of note, some implications of the findings in Experiment 8 and 9 may lead to different observations. For example, the additional depth information involved in the two experiments, the different experimental environments (e.g., the different luminance of the environment and the visual stimuli, the visual distance, the use of the 3D glasses and so on) may lead to different observations. In order to rule out these effects, Experiment 10 was designed and conducted in the same cabin as used Experiment 8.

4.5. Experiment 10

4.5.1. Materials and Methods

The method was essentially the same as in Experiment 8, and both the configural context and the pseudo cuboid were presented on the same depth plane; no depth information was involved in this experiment. It should be note that, in this experiment, the pseudo cuboid was controlled as a background task-irrelevant information (similar as in Experiment 9) to further clarify the foreground-background segmentation effect in contextual cueing. In order to achieve this, the search items were not constrained to the edges of the cuboid shape, rather they were located anywhere else (see Figure 4.1 C, left panel, the 78 possible item locations inside the whole 13.2° x 13.2° square area). With this arrangement, the association between configural context and the “up-tilted” cuboid is weak, thereupon the pseudo cuboid was not associated to the locations of the items can be considered as background task-irrelevant information. It is important to point out that in the first transfer session, in which the cuboid was rotated 90° (“down-tilted” cuboid), a part of visual search items (around 15.4% of the overall items) do appear on the edges of the “down-tilted” cuboid, which may bring the background cuboid to participant’s focus of attention to some degree. Same as in Experiment 8 and 9, only participants who showed positive contextual cueing effect during the training session was selected for further analysis (10 participants with mean age of 26.5 years old were selected, 7 of them are females).

4.5.2. Results

Search task. Similar to the analysis used in Experiment 8, the trials were grouped into epochs 1-4 for the training session, and epochs 5 and 6 for the transfer session. Trials with erroneous responses or reaction times (RTs) that were outside the range of 0.2s and 3 standard deviations were excluded for further analysis. Both the error rates and the ratio of outliers were low (mean error rates: 1.0%; outliers: 2.52%). The error rates were comparable across all conditions: context, $F(1, 9) = 0.02$, $p = 0.89$, $\eta_p^2 = 0.002$, epoch, $F(2.52, 22.67) = 1.83$, $p = 0.13$, $\eta_p^2 = 0.17$, and interaction, $F(5, 45) = 0.59$, $p = 0.71$, $\eta_p^2 = 0.061$), showing no improvement in accuracy by training. The overall results are presented in Figure 4.2 C.

A repeated-measures ANOVA with context (old vs. new) and epoch (1-4) as factors was applied to mean RTs and revealed significant differences of epoch, $F(1.49, 13.41) = 13.31$, $p < 0.001$, $\eta_p^2 = 0.60$, with 316 ms faster in epoch 4 compared to epoch 1; as well as of context, $F(1, 9) = 34.39$, $p < 0.001$, $\eta_p^2 = 0.79$, with 292 ms faster for the old display compared to the new display. The interaction was also significant, $F(3, 24) = 7.58$, $p < 0.01$,

$\eta_p^2 = 0.46$. Further paired sample t-test revealed significant contextual benefit of epoch 2-4 ($p < 0.01$) but only marginally significant of epoch 1 ($p = 0.05$), indicating both contextual cueing effect and learning effect was manipulated in the training session. One-way repeated measures ANOVA analysis in the transfer sessions revealed significant contextual cueing transfer effect in both Epoch 5 with “down-tilted” cuboid, $F(1,9) = 11.02$, $p < 0.01$, $\eta_p^2 = 0.55$, mean effect of 321 ms, as well as in Epoch 6 without cuboid, $F(1,9) = 8.39$, $p < 0.05$, $\eta_p^2 = 0.48$, mean effect of 137 ms, suggesting the learned contextual cueing facilitation during the training session with a background cuboid can be transferred when the cuboid was varied or removed. These results stark contrast to the findings obtained in Experiment 8, where a foreground cuboid was used.

Recognition results. All the trials were responded within 20 s. The mean hit rate (58.75%) was significantly higher than the false alarm rate (43.75%), $F(1, 9) = 7.36$, $p < 0.05$, $\eta_p^2 = 0.45$. Further analysis revealed that 5 out of 10 participants were able to recognize the old displays by showing at least 25% higher hit rate than the false alarm rate. These results suggested that participants were able to remember the old configurations. The relative better recognition performance in the current experiment compared to previous experiments may due to the grouping effect of the background cuboid. In this experiment, the visual search items were easily separated by the cuboid into 5 sub-areas (e.g., inside or outside the cuboid), with a mean of 2-3 items in each area; while in Experiment 8 and 9, the visual items were on the edges of the cuboid and are difficult to be grouped into sub-areas. The relative small number of visual search items in each area are can boost a better explicit memory performance. Nevertheless, explicit memory in the current experiment does not help participants to maintain the well-established contextual cueing when the cuboid varied or removed. These observations confirm our previous study (Shi et al., 2013) that suggest recognition and visual search may involve different memory process, with the former recruiting more complex information-matching processes that do not facilitate the search processing.

4.5.3. Discussion

In Experiment 10, we presented visual search items together with a background cuboid on the same depth plane to examine whether the background cuboid can be involved in contextual learning. The results showed the learned contextual facilitation during training session can be transferred to the display with a varied or removed cuboid. Providing that the contextual cueing was manifested in both training and transfer sessions, the background cuboid, which was merely presented together with the repeated configural context, was probably not encoded into spatial contextual memory. Considering the same experimental environments (i.e., same luminance level of the environments and visual stimuli, visual distance, no depth information) in Experiment 8 and 10, our results confirm that the environmental differences are not the reason which leads to different observations in the first two experiments. Instead, the findings in Experiment 10 further confirmed our hypothesis that foreground-background segmentation influences contextual cueing effect, proving a foreground task-relevant context was learned together with the configural context, but a background context was ignored during contextual learning.

4.6. General Discussion

In the three experiments conducted in the current study, participants searched for a target on the display consisting of both visual search items and an additional item-independent pseudo 3D cuboid. The aim of the manipulation was to investigate the interaction effects (if any) of foreground-background segmentation and contextual cueing. In order to assess different contributions of foreground/background context information in contextual guided visual search, the properties (foreground versus background) of a pseudo 3D cuboid were manipulated in three experiments. In the training session of Experiments 8 and 9, visual search items were allocated on the edges of the pseudo cuboid but not to the empty space, whereas the reverse was true for Experiment 10, that is, the items were allocated off the edges in the training session. Of note, the pseudo cuboid was presented on the back of the fixation plane (contains visual search items) in Experiment 9 with 3D disparity vision. On this ground, the pseudo cuboid served as foreground information in Experiment 8 but background information in Experiment 9 and 10. In the subsequent transfer sessions of all three experiments, the pseudo cuboid was either rotated 90°, or removed, without any changes of the configural context. The results showed that contextual cueing effects were transferred from the training to the transfer sessions, only when the cuboid was presented as background context in Experiment 9 and 10, but not when the cuboid served as foreground context in Experiment 8, suggesting foreground-background segmentation interferes with the contextual cueing. Further recognition test revealed that participants explicitly recognized the repeated configural context from the novel one in Experiment 10 but not 8 and 9. However, this explicit memory was not correlated to contextual cueing transfer effect when the cuboid varied, suggesting recognition and visual search may involve different memory process.

Our finding that the well-established contextual cueing was maintained when the background cuboid but not the foreground cuboid varied suggests that the effect of the item-independent shape information in contextual cueing depends on its foreground/background properties. When the visual search items were presented on the edges of the cuboid in experiment 1, different items were connected by the 9 white lines that formed the edges of the cuboid, therefore providing a uniform connectedness of the visual search items. As proposed by previous studies (Han et al., 1999; Palmer, 1992; Palmer & Rock, 1994), the uniform connectedness provided an effective way of perceptual grouping, thus, the connected visual search items in Experiment 8 can be grouped together and formed a unique “cuboid-like” object. On this ground, we suggest that both the visual search items and the cuboid were perceived as foreground context therefore attracting participants’ attention after the preattentive segmentation stage (Treisman & Gelade, 1980; Wolfe, 1994b, 2003a) and encoded in contextual memory. As a result, the spatial configuration was learned with reference to the foreground cuboid in Experiment 8, although the cuboid did not serve as spatial cue and could not predict the target’s location. Nevertheless, when presenting the visual search items hinter the item plane with disparity vision in Experiment 9, the frontal item context can be more salient compared to the hinter cuboid therefore restricting attention to the frontal plane and sequestering attentive processing of the hinter plane. The saliency effect generated by objects lying in the front plane, thus attracting more attentional resources than those lying in the hinter plane was also observed by previous studies (Mazza et al., 2005; Nakayama, Shimojo, & Silvermann, 1989). In the Experiment 10, the spatial relationship between visual search items and the cuboid was weak even although both of types of the visual information were presented with the same depth plane. As a result, the items could be easily segregated from the cuboid therefore attracted participants’ attention. In this case, the learned contextual facilitation mainly depended on the configural context while the variation of the background cuboid in the transfer session didn’t impede contextual cueing. Based on these observations, we conclude that, the foreground-background segmentation that occurs during preattentive stage affects participant’s deployment of attention, herein leading participants to deploy more attentional resources to the foreground context than to the background information during the visual search task.

The finding that contextual learning and retrieving occurs after image segmentation is indeed an interesting pheromone. Whereas on one hand, previous studies on contextual cueing proposed that implicitly learned spatial memory can guide participants’ attention to possible targets location (e.g., Annac et al., 2013; Assumpção, Shi, Zang, Müller, & Geyer, 2015; Brady & Chun, 2007; Chun & Jiang, 1999), boosting a better search performance of the repeated displays, on the other hand, other studies (Endo & Takeda, 2004; Jiang & Chun, 2001; Jiang & Leung, 2005) proposed that attention modulate contextual retrieving. However, none of them have investigated at which selective attention (preattentive v.s. attentive) stage contextual learning and retrieving occurs. Due to the reason that the filtering-out item-independent background cuboid by the image segmentation processing stage was not encoded in contextual learning, and the image segmentation usually occurs during preattentive stage (e.g., Driver et al., 2001a; Nakayama & Silvermann, 1986), we can speculate that contextual learning may occurs after preattentive stage. In other words contextual learning occurs during attentive stage. This speculation was also hinted by previous studies (Geringswald et al., 2012; Geringswald et al., 2013a; Zang et al., 2013). For example contextual cueing was observed to developed but not retrieved under the tunnel vision manipulation that participants can only see the configuration falling inside a limited area around their fixations (Zang et al., 2013). Therefore, the missing global context makes preattentive processes such as segmentation or grouping obsolete for the task, thus advancing to attentive processes in order to learn the local context around fixations. These successive samples of the local information obtained by para-/foveal vision were then constructed to an overall representation (e.g., De Graef, 2007; Deubel et al., 2010) thus boosting contextual facilitation. Interestingly, when limiting or removing para-/foveal local information which may impede the attentive processing, the repeated spatial context could not be learned. For example, contextual cueing never developed for repeated displays with simulated foveal scotoma or fovea degeneration (Geringswald et al., 2012; Geringswald et al., 2013a). All in all, although more direct evidences should be considered in future studies, it is quite possible that contextual cueing occurs during attentive but not preattentive processing.

Our findings that the foreground but not background visual information was encoded during spatial context learning may provide a unified account for the controversial findings in previous studies (Brooks et al., 2010; Kunar, John, et al., 2013; Rosenbaum & Jiang, 2013) which investigated the interactions between predictive configural context and predictive item-independent scene context. For example, Kunar, John, et al. (2013) has demonstrated that when the contextual learning was manifested by repeating both the background color and the configural search array, preserving either background or configural context was enough to maintain contextual cueing. Nevertheless, contextual cueing was largely preserved when the configural context was maintained (180ms), and weakly preserved when only the background context was available (77ms). In contrast to Kunar, John, et al. (2013)'s findings, another study from Brooks et al. (2010) observed a joint learning effect of a natural scene context and the configural context. The learned contextual cueing of displays with both scene and configural predictive context cannot be retrieved when one type of the contexts was missing. A recent study from Rosenbaum and Jiang (2013) makes the view of the contextual interaction even more complex, by observing the learned contextual cueing from both configural predictive and scene predictive displays could only be transferred to the scene predictive (configuration varied) display, but not to the configural predictive (scene varied) display. These studies, although with similar experimental settings, observed contradictory results, and these contradictions have not been well explained. Fortunately, our findings that foreground information is more important in contextual cueing may provide a possible explanation. In the study of Kunar, John, et al. (2013), the unique color information, although predictive to the target's location, contained little spatial relations to the search array. Thus the color could be segregated as the background context. In this case, it is reasonable that the configural learning served as the major source that driving contextual cueing. In Brooks et al. (2010)'s experiments, the visual search array (i.e., target and distractors) was presented on the center area of a real nature scene on the peripheral area (e.g., a room scene with chairs and tables) of a display. Considering the real nature scene normally contains salient objects (e.g., a chair) which may attract participants' attention, both the configural context and the nature scene could be segregated as foreground context to some extent. In this case, repeated spatial context could be learned depending on both types of the context information. Therefore, the learned contextual cueing facilitation providing both types of context cannot be preserved when one type of the context was missing. Different to Brooks et al. (2010), the predicted natural scene of the display in Rosenbaum and Jiang (2013)'s study was spread over the display (with a size of 20°×20°). The relatively small visual search items (i.e., one "T" and eleven "L"s with the size of 0.56°×0.56°) were sparsely embedded over the natural scene. In this way, the configural context was much less salient than the predicted natural scene. Therefore the contextual cueing was learned mainly based on the scene context which was segregated as foreground information but not dependent on the configural context. To conclude, our findings provide direct evidence for the effect of foreground-background association in contextual learning and retrieving, and valuable evidences to explain the contradictory findings of the previous studies on the interactions of configural context and item-independent features.

In summary, the present findings suggest that the influence of item-independent pseudo 3D shape on contextual cueing depends on it foreground/background properties on the visual search display. When the pseudo 3d shape is segmented together with the configural context, resulting in a putatively foreground perception, any changes or removal of the pseudo 3D shape may abolish the learned contextual cueing. However, when the pseudo 3D shape serves as background task-irrelevant information, the contextual cueing is less influenced by the changes of the background context. Our findings highlight the role of the foreground and background contexts in the contextual learning, and point out the possible processing stage of contextual learning (attentive stage). The foreground/background segmentation between different types of context may guide participants’ attention to particular information on the visual display during contextual learning, thus modify the roles of different contexts in contextual memory. It should be noted, however, in our study, we only applied a pseudo 3D cuboid as the item-independent context. To generalize our conclusion, it is necessary to use other background contexts in future studies. Nevertheless, our study provides promising unification of the influence of foreground-background segmentation in the contextual learning.

Chapter 5. Deutsche Zusammenfassung

Visuelle Suche, wie z.B., die Suche nach dem Hefter auf einem Schreibtisch oder die Suche nach der Lieblings-Schokolade in einem Supermarkt, ist eine der gewöhnlichsten Aufgaben im Alltag, die normalerweise ein aktives Absuchen der visuellen Umgebung nach einem bestimmten Ziel beinhaltet. Die Fähigkeit, ein Objekt in einem komplexen Umfeld effizient zu lokalisieren wurde im letzten halben Jahrhundert ausgiebig untersucht und viele Faktoren, die in der „bottom-up" und „top-down"-Verarbeitung von visueller Suche beinhaltet sind, konnten identifiziert werden (Wolfe, 2003b). Um effektiv mit der Außenwelt interagieren zu können, verwenden wir oft relevante Informationen während wir irrelevante Informationen ignorieren, um unsere Aufmerksamkeit auf das Ziel zu lenken (Wolfe, 1994a). Eine Art solcher relevanter Information ist die Unveränderlichkeit des visuellen Inputs (Chun, 2000). In den meisten Fällen ist unsere visuelle Außenwelt stark strukturiert und über die Zeit hinweg stabil. Folglich ist es für Wahrnehmungsprozesse sehr vorteilhaft, empfindlich auf die unveränderte Struktur zu reagieren, da es die Vorhersehbarkeit der Außenwelt steigert (Gibson, 1969). Beispielsweise würde die Verwendung unveränderter Hinweise, wie etwa die gewöhnliche Position eines Hefters, die sich auf dem Schreibtisch aber nicht unter dem Schreibtisch befindet, die Suche nach dem Hefter erleichtern.

Die Erleichterung durch den unveränderten Kontexts bei der visuellen Suche, die im „Contextual Cueing"-Paradigma erfasst wird, hat großes Interesse in der Gemeinschaft der visuellen Suche geweckt (z.B., Chun, 2000; Chun & Jiang, 1998; Conci & Müller, 2012; Geringswald et al., 2012; Geringswald, Herbik, Hoffmann, & Pollmann, 2013b; Manginelli, Langer, Klose, & Pollmann, 2013; McDonnell, Mills, McCuller, & Dodd, 2014). In einem klassischen „Contextual Cueing"-Paradigma werden Teilnehmer angewiesen, unter einer Reihe von Distraktoren einen Zielreiz auf dem visuellen Suchbildschirm zu suchen. Ohne die

Teilnehmer zu informieren, wird die Hälfte der Displays (Anordnung von Zielreiz und Distraktoren auf dem Bildschirm) während des Experiments wiederholt dargeboten, während die andere Hälfte der Displays immer neu generiert und im Laufe des Experiments nicht wiederholt wird. Die Teilnehmer reagieren in der Regel schneller auf bereits gezeigte Displays im Vergleich zu den neu generierten Displays („Contextual Cueing“-Effekt), was darauf hindeutet, dass der unveränderte Kontext tatsächlich die Suchleistung verbessern kann (Chun & Jiang, 1998), 1998).

Da der unveränderte Kontext durch verschiedene Arten von visuellen Strukturen zustande kommen kann, wie z.B., die lokale Anordnung von Suchelementen, die globale Struktur oder Hintergrundeigenschaften, ist es wichtig, herauszufinden, welche Rolle diese verschiedenen Typen von Kontexten spielen, um den Mechanismus des „Contextual Cueing“-Effekts zu verstehen. Eine Reihe von aktuellen Studien haben die unterschiedliche Rollen des lokalen, des globalen sowie des Hintergrundkontexts im „Contextual Cueing“-Effekt untersucht: So haben einige Studien eine große Bedeutung der lokalen Unveränderlichkeit im kontextuellen Lernen gefunden. So wurde zum Beispiel gezeigt, dass eine Wiederholung von drei bis vier lokalen Elementen in der Nähe des Zielorts (lokaler Kontext) genügt, um einen „Contextual Cueing“-Effekt zu produzieren (Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005). Auf der anderen Seite zeigten weitere Studien (Brockmole et al., 2006; Brockmole & Henderson, 2006a) mit Hilfe von Natur-Szenerien als Reize für die visuelle Suche, dass der globale unveränderte Kontext wichtig ist, um den gelernten „Contextual Cueing“-Effekt zu erhalten. Es wird gezeigt, dass ein gut etablierter „Contextual Cueing“-Effekt nur übertragen werden kann, wenn der globale Kontext unverändert bleibt und dieser unabhängig von der Änderung des lokalen Kontexts ist, aber nicht umgekehrt. Außerdem haben Kunar et al. (2006) festgestellt, dass das Hintergrundmuster auch als Prädiktor zur Erleichterung der visuellen Suche verwendet werden kann. Obwohl diese Studien wertvolle

Ergebnisse darlegen, die uns helfen, die Rollen von verschiedenen Kontexten zu verstehen, sind noch eine Reihe interessanter Fragen offen: Da sich in früheren Studien die lokalen und globalen Kontexte nicht vollständig gegenseitig ausschließen, ist es noch nicht klar, ob ein rein lokal fovealer Kontext ohne globalen peripheren Kontext ausreicht, um einen „Contextual Cueing"-Effekt zu induzieren. Außerdem ist auch nicht bekannt, wie die globalen und lokalen Kontexte während des kontextuellen Lernens und der Wiederauffindung von Reizen miteinander interagieren. Darüber hinaus wird die Interaktion von Hintergrund- und Vordergrundkontexten im kontextuellen Lernen ebenfalls diskutiert.

Aus diesem Grund hat die vorliegende Arbeit, welche aus drei Studien besteht, das Ziel, die oben genannten offenen Fragen in Bezug auf die unterschiedlichen Rollen des lokalen und globalen Kontexts sowie des Hintergrundkontexts zu untersuchen. Durch die Kombination einer durch Augenbewegungen gesteuerten Technik (Duchowski, Cournia, & Murphy, 2004; Loschky & McConkie, 2002) und des klassischen „Contextual Cueing"-Paradigmas untersucht Studie 1 die Wechselwirkungen der lokalen und globalen Kontexte im kontextuellen Lernen und beim Wiederauffinden von Reizen. Studie 2 untersucht die Rolle der lokalen invarianten Informationen bei der Übertragung des gelernten „Contextual Cueing"-Effekts nach Änderung des globalen Kontexts. Zuletzt untersucht die 3. Studie den Einfluss der Assoziationsstärke zwischen dem Vorder- und Hintergrund-Kontext in einer kontextuell geführten visuellen Suche.

5.1. Die Interaktion zwischen der fovealen lokalen und peripheren globalen Information im kontextuellen Lernen und im Gedächtnisabruf

Die erste Studie konzentrierte sich auf die Interaktion der fovealen lokalen und peripheren globalen Kontexte im kontextuellen Lernen und im Gedächtnisabruf. Insbesondere untersuchten wir, ob reine foveale lokale Information, ohne periphere globale Information, ausreichend ist, um einen „Contextual Cueing"-Effekt zu generieren. Mit anderen Worten, ist es notwendig, ein gewisses Maß an globalem Kontext zu haben, um einen „Contextual Cueing"-Effekt zu entwickeln? Um wirklich keine periphere Information bei der Präsentation der Stimuli abzubilden, setzten wir eine augenbewegungsgesteuerte Technik zur Einschränkung der Sicht in der Trainingseinheit ein (Loschky & McConkie, 2002; Loschky & McConkie, 2000; Parkhurst et al., 2000). Diese Technik beschränkt zu jedem Zeitpunkt die Sicht auf das Gebiet um die Fixierung. Die Größe des augenbewegungsgesteuerten eingeschränkten Sichtkreises wurde in verschiedenen Experimenten auf 8° und 12° festgesetzt. Die Teilnehmer mussten den Bildschirm aktiv erkunden, um den Zielreiz unter einer Menge von Distraktoren herauszufinden. Um zu überprüfen, ob der unter der augenbewegungsgesteuerten eingeschränkten Sicht gelernte „Contextual Cueing"-Effekt, wenn er überhaupt auftritt, auch auf einen Bildschirm mit uneingeschränkter Sicht übertragen wird oder nicht, wurde in den folgenden Transfersitzungen die eingeschränkte Sicht entfernt und die gesamte Anzeige präsentiert. Die Ergebnisse zeigten keinen signifikanten „Contextual Cueing"-Effekt während des Trainingsprozesses, wenn der foveale lokale Sichtkreis klein war (8°, mit einem Mittelwert von 2,09 von 12 Elementen), aber einen deutlichen „Contextual Cueing"-Effekt wenn der Sichtkreis groß war (12°, mit einem Mittelwert von 4,69 von 12 Elementen). Diese Ergebnisse zeigen, dass der „Contextual Cueing"-Effekt von der Größe der sichtbaren räumlichen Konfiguration abhängt, und dass das Fehlen peripherer visueller Information in einer relativen sichtbaren Anzeigengröße (8° Blickwinkel) den Gedächtnisabruf von unverändertem Kontext verhindert. Durch Hinzufügen einer kurzen Vorschau (150 ms) der gesamten räumlichen Konfiguration vor der augenbewegungsgesteuerten eingeschränkten Sicht (8°) erholte sich jedoch der „Contextual Cueing"-Effekt. Entscheidend ist, dass alle Experimente signifikante „Contextual Cueing"-Effekte in den Transfersitzungen zeigten, wenn die gesamte Anzeige sichtbar war. Diese Ergebnisse deuten darauf hin, dass der räumlich unveränderte Kontext gelernt und basierend auf ausschließlich begrenzter fovealer Information konstruiert werden kann, aber zum Gedächtnisabruf des gelernten Kontexts einige periphere Informationen oder eine kurze globale Vorschau der gesamten Anzeige erforderlich sind.

Außerdem ergab eine Analyse des okulomotorischen Verhaltens eine signifikante Verringerung der Zahl der Fixationen bei der alten Anzeige im Vergleich zur neuen Anzeige, was mit früheren Studien konsistent ist (Peterson & Kramer, 2001; Tseng & Li, 2004; Zhao et al., 2012). Das Ergebnis zeigt, dass die wiederholte kontextuelle Konfiguration als Leitfaden für die Suche nach dem Zielreiz verwendet werden kann. Darüber hinaus wurde der „Contextual Cueing"-Effekt mit einer erhöhten mittleren Fixierungsdauer verbunden, was wahrscheinlich eine verbesserte Planung des Wegs der Sakkade zum Zielort reflektiert (Zou et al., 2012).

Abschließend ist zu sagen, dass die Anwendung einer augenbewegungsgesteuerten Anzeigenvariation in einem klassischen „Contextual Cueing"-Paradigma uns in Studie 1 gestattete, die globalen und lokalen Kontexte sauber zu trennen. Die Ergebnisse zeigen eindeutig, dass räumlich unveränderte Kontexte basierend auf lediglich 2-3 lokal verfügbaren Elementen gelernt werden können, aber die gelernten Kontexte nur wieder abgerufen werden können, wenn einige weitere periphere (globale) Informationen verfügbar sind.

5.2. Die Rolle der lokalen unveränderten Information bei der Übertragung des „Contextual Cueing“-Effekts nach Änderung des Anzeigemodus

In Studie 1 fanden wir heraus, dass der lokale Kontext eine wichtige Rolle beim kontextuellen Lernen spielt, während der globale Kontext als Hilfe gebraucht wird, um sich an den erlernten Kontext zu erinnern. Zahlreiche aktuelle Studien (Brady & Chun, 2007; Olson & Chun, 2001; Song & Jiang, 2005) haben herausgefunden, dass sich kontextuelles Lernen zeigt, wenn der lokale unveränderte Kontext bestehen bleibt (z.B., 3-4 visuelle Suchitems in der Nähe des Zielreizes konstant halten), der globale Kontext jedoch verändert wird (die restliche visuelle Information). Aber vorherige Studien haben eine konstante Displaymatrix verwendet, bei der die Orientierung des Suchdisplays unverändert blieb. Obwohl sich der Displaymodus für den normalen Monitor kaum verändert (z.B., Querformat), kann er sich bei aktuellen, auf Berührung basierenden Mobilfunkgeräten häufig ändern (z.B. der Wechsel des Displays zwischen dem Querformat und dem Hochformat eines iPads). Deshalb ist es wichtig, herauszufinden, wie sich ein Wechsel des Displaymodus auf das Suchverhalten auswirkt, so wie z.B., das Ausfindigmachen eines Zielsymbols bei einem Mobilfunkgerät.

Um herauszufinden, ob der bereits erlernte Kontext erhalten bleiben kann, wenn der globale Anzeigemodus vom Querformat zum Hochformat wechselt, hat Studie 2 den Einfluss von 4 verschiedenen Methoden, in denen die Icons jeweils neu sortiert waren, auf das Suchverhalten verglichen: 1) „unveränderte Reihenfolge“: die von Mobilfunkgeräten verwendete einzig verfügbare Neuanordnung der Items: Die Position der Icons in einem Anzeigemodus wird auf den anderen Anzeigemodus übertragen, indem einfach die Anordnung der Icons beibehalten wird von der linken oberen Ecke bis zur rechten unteren

Ecke. Dabei ist es wichtig zu erwähnen, dass bei der „unveränderten Reihenfolge“ meist die lokalen Beziehungen der Items zueinander zerstört werden, besonders wenn der ganze Display ein Rechteck ist (was bei den meisten Mobilfunkgeräten der Fall ist.) 2) „globale Drehung“: Vergleichbar mit der Drehung eines Objekts in unserer physikalischen Welt, dreht sich der gesamte Display um 90° im Uhrzeigersinn während die individuelle Ausrichtung der Icons direkt nach der Neuanordnung beibehalten wird. 2) „lokal unverändert“: fünf lokale Bereiche der gesamten Konfiguration bleiben nach der Neuanordnung der Icons erhalten. 4) „zentral unverändert“: Das zentrum des Displays bleibt nach der Neuanordnung der Icons unverändert. Sowohl die Methoden „lokal unverändert“ als auch „zentral unverändert“ zielen auf einen maximal konstanten lokalen Kontext nach der Neuanordnung ab. Wir verwendeten das klassische kontextuelle Paradigma für diese 4 Neuanordnungs-Methoden, um herauszufinden, bei welcher Methode der Neuanordnung der erlernte unveränderte Kontext in einem Anzeigemodus auf den anderen Anzeigemodus übertragen werden kann. Der visuelle Display, der in Studie 2 verwendet wurde, setzte sich aus 24 alltäglich verwendeten Icons (z.B., das Skype-Icon) zusammen, die entweder im Querformat (ein 4x6-Format mit 4 Icons pro Spalte und 6 Icons pro Reihe) oder im Hochformat (ein 6x4-Format mit 6 Icons pro Spalte und 4 Icons pro Reihe, wie in Figur 3.2). Die Teilnehmer wurden instruiert, nach einem Zielicon zu suchen, das mit dem Buchstaben „T“ überspielt wurde. Die visuelle Suchaufgabe der vorliegenden 4 Experimente in Studie 2 beinhaltete 2 Sitzungen: Das Training und die Testsitzung. Während des Trainings wurden 12 unveränderte Displays im Querformat zufälligerweise generiert und wiederholt präsentiert. In der anschließenden Testsitzung haben wir in der Hälfte der Durchgänge neue Displays eingeführt, um zu vergleichen, ob ein kontextueller Effekt im selben Anzeigemodus zustandekommen kann und ob der kontextuelle Effekt auf den anderen Anzeigemodus übertragen werden kann, basierend auf den eben genannten 4 Methoden der Neuanordnung der Icons. Die Ergebnisse enthüllten einen robusten

„Contextual Cueing"-Effekt im Training für alle 4 Experimente und die erlernte kontextuelle Erleichterung blieb auch nach zwei dazwischenliegenden Blocks mit dem Hochformat-Modus erhalten. Interessanterweise wurde der erlernte „Contextual Cueing"-Effekt für das Querformat nur dann erfolgreich auf das Hochformat übertragen, wenn die Icons auf dem Display durch die „lokal unveränderte" und die „zentral unveränderte" Methode neu angeordnet wurden, aber nicht bei der „unveränderten Reihenfolge" und der „globalen Drehung" Methode. Diese Ergebnisse lassen darauf schließen, dass die klassische „unveränderte Reihenfolge"-Anordnungsmethode, die in aktuellen Mobilfunkgeräten verwendet wird, nicht die optimale Methode ist, um das Suchverhalten des Gebrauchers zu verbessern. Zudem ist die „globale Drehung" Methode, obwohl sie für Benutzer intuitiv erscheint, nicht von Vorteil für das Suchverhalten des Benutzers nach dem Wechsel des Anzeigemodus, da wahrscheinlich zusätzliche Prozesse in der mentalen Drehung für den gedrehten Anzeigemodus benötigt werden (Böckler et al., 2011; Borst et al., 2011; Ionta & Blanke, 2009; Shepard & Metzler, 1971; Shomstein & Yantis, 2004). Hervorzuheben ist, dass der erlernte Kontext bei der „lokal unveränderten" und „zentral unveränderten" Methode beibehalten werden konnte, da das Beibehalten von lokal regionalen Konfigurationen und der topologischen Beziehung zwischen diesen Regionen der Schlüssel zum Erhalt des erlernten räumlichen Kontext beim Wechsel des Anzeigemodus ist.

5.3. Einflüsse von Informationen des Vorder- und Hintergrunds im räumlichen Kontextlernen und Kontextabruf

Die drei Experimente (Experimente 8, 9 und 10) des dritten Teils der Dissertation konzentrieren sich auf die Interaktion zwischen Vordergrund und Hintergrund in der kontextuell geführten visuellen Suche. Genauer gesagt konzentrierten wir uns auf die Frage, wie die unterschiedlichen Kontexte zwischen Hintergrund- und Vordergrundinformationen den „Contextual Cueing“-Effekt beeinflussen. Ein klassisches „Contextual Cueing“-Paradigma wurde in drei Experimenten herangezogen. Zusätzlich zu dem klassischen „Contextual Cueing“-Paradigma wurde eine item-unabhängige geometrische Form (d.h. ein Rahmen in Form eines Quaders) unter der Item-Anordnung (die 1 „T“ und 11 „L“s beinhaltet) angezeigt. Wichtig ist, dass die Eigenschaften der item-unabhängigen Informationen durch Änderung der Positionen und der Tiefen zwischen der Item-Anordnung und der Quader, während der Trainingsphase auf verschiedene Arten variiert wurde. Die geometrische Form wurde als Vordergrund der Trainingsphase des Experiments 8 aber als Hintergrund der Experimente 9 und 10 eingestellt. In der anschließenden Transfersitzung wurde der Quader entweder um 90° gedreht oder ganz entfernt, um zu untersuchen, ob das „Contextual Cueing“ durch die Änderung der Quader-Form beeinflusst wird. Die Ergebnisse der Trainingssitzungen bestätigten einen signifikanten „Contextual Cueing“-Effekt in allen Experimenten. Während der anschließenden Transfersitzungen zeigte sich, dass der erlernte „Contextual Cueing“-Effekt zwar in der Bedingung zerstört wurde, in der die Änderung der Vordergrund-Quader stattfand (Experiment 8), jedoch nicht in derjenigen, in der die Änderung der Hintergrund-Quader (Experimente 9 und 10) stattfand. Diese Ergebnisse lassen darauf schließen, dass der Quader und die visuellen Items im Vordergrund während des kontextuellen Lernens zusammen erworben werden, wenn der Quader als Vordergrund Informationen präsentiert wurde. Als Ergebnis lässt sich festhalten, dass der erlernte „Contextual Cueing“-Effekt bei einem Wechsel des Quader -Kontexts empfindlich ist und dass jede geringe Änderung des Hintergrunds den erlernten „Contextual Cueing“-Effekt zerstören kann. Im Gegensatz dazu kann Anordnung des Vordergrund-Kontexts unabhängig vom Hintergrund während der Trainingsphase erlernt werden, und der erlernte „Contextual Cueing“-Effect kann auf Displays mit einem anderen Hintergrund übertragen werden.

Zusammenfassend lässt sich sagen, dass die Ergebnisse der dritten Studie einen direkten Beweis für den Effekt der Assoziation zwischen Hintergrund und Vordergrund im kontextuellen Lernen und Wiederauffinden liefern. Die Quaderinformation wird während dem kontextuellen Lernen quasi ignoriert, wenn die Quaderinformation auf verschiedene Tiefenebene im Vergleich zu der Item-Anordnung präsentiert wird (z.B. der Quader, der in Experiment 9 verwendet wurde), oder wenn sie nicht stark mit den Items im Vordergrund assoziiert wird (z.B. der Quader, der in Experiment 10 verwendet wurde). Wenn die Quaderinformation jedoch stark mit dem Kontext im Vordergrund assoziiert wird (z.B., der Quader, der in Experiment 8 verwendet wurde), wird der Quaderkontext meist in das kontextuelle Lernen bei der Anordnung der Reize im Vordergrund integriert. Deshalb kann der bereits erlernte „Contextual Cueing“-Effekt bei verändertem Quader zerstört werden.

5.4. Bewusstsein und okulomotorische Korrelationen in der kontextuell geleiteten visuellen Suche

In allen vorliegenden Studien wurde das Bewusstsein über das erlernte kontextuelle Gedächtnis erforscht. Es wurde sowohl implizites als auch explizites Lernen gefunden: 1) In der ersten Studie konnten die Teilnehmer wie in vielen anderen Studien zum „Contextual Cueing“ (z.B., Chua & Chun, 2003; Cleeremans, Destrebecqz, & Boyer, 1998; Jiang, Won, et al., 2014; McDonnell et al., 2014) nicht explizit die wiederholten bzw. neuen Displays identifizieren, wenn die Displays nur in einem durch die Blickbewegung eingeschränkten Bereich sichtbar waren (Experimente 1 und 2). Jedoch führte eine kurze, zusätzliche Vorschau des ganzen Displays vor jedem Durchgang zu explizitem Lernen (Experiment 3). Wir vermuteten, dass die kurze Vorschau auf das gesamte Display vor der durch die Blickbewegung begrenzten Sicht zusätzliche Aufmerksamkeit auf die räumliche Konfiguration des gesamten Displays lenkt. Darum kann man die Konfigurationen explizit lernen. Dennoch sollte beachtet werden, dass das explizite Lernen nicht mit der Stärke des Kontexteffekts korreliert, was darauf hindeutet, dass die Verwendung von explizitem Lernen nicht unbedingt einen zusätzlichen Beitrag zum Kontexteffekt bringt. 2) In Studie 2 erkannten Teilnehmer nicht nur die alten Landschaftdisplays, sondern auch die neu zugeordneten Porträtdisplays in allen vier Experimenten (4-7), möglicherweise aufgrund der umfangreichen Eigenschaften, die durch die verschiedenen Symbole geboten werden. Sie bestehen aus Reizen für visuelle Suche. Aber ähnlich wie in Studie 1 korreliert das explizite Lernen nicht mit dem Kontexteffekt, was nur in 2 von 4 Display-Methoden herausgefunden wurde. 3) In Studie 3 wurde das traditionelle implizite Lernen für das erste zwei Experiment beobachtet. Wir fanden jedoch explizites Lernen, wenn visuelle Suchelemente zufällig auf einem Hintergrundquader dargestellt wurden. Wir glauben, dass das explizite Lernen teilweise aus der Segmentierungswirkung resultiert, die von der Hintergrundform verursacht wird. Das heißt, die Suchelemente wurden durch die drei Flächen des Hintergrundquaders in mehrere Unterbereiche eingeteilt, deshalb erhöht es die Wahrscheinlichkeit, sich an einen oder mehrere Unterbereiche anstatt an das gesamte Display zu erinnern. Das explizite Lernen von solchen Teilbereichen wäre der wichtigste Faktor für die explizite Erinnerung der Konfigurationen in den Teilbereichen. Auch hier korrelierte das explizite Gedächtnis der Konfigurationen nicht mit dem Kontexteffekt, was die Ergebnisse aus den Studien 1 und 2 bekräftigt. Zusammengefasst können die Wiedererkennung und die visuelle Suche verschiedene Gedächtnisprozesse beinhalten und das Kontextwissen kann durch implizite und explizite Lernprozesse erworben werden.

Zusätzlich zu den Analysen der manuellen Reaktionen und des Bewusstseins für kontextuelles Lernen wurden auch die Augenbewegungen der Teilnehmer für Studien 1 aufgenommen und analysiert. Die okulomotorischen Ergebnisse zeigten, dass der Kontexteffekt mit einer Reduktion der Anzahl der Sakkaden sowie mit einer Erhöhung der

Fixationsdauer gekoppelt ist. Die gekoppelte Beziehung zwischen der Anzahl der Sakkaden und Kontexteffekt wurde auch in früheren Studien beobachtet (z.B., Manginelli & Pollmann, 2009; Ogawa & Watanabe, 2010; Zhao et al., 2012), was darauf hinweist, dass der alte Kontext die Aufmerksamkeit der Teilnehmer auf die Position des Zielreizs lenkt und einen effizienteren Suchpfad erzeugen kann. Die erhöhte Fixationsdauer in der wiederholten Konfiguration stellt im Vergleich zu der neuen Konfiguration während der kontextuell geführten visuellen Suche eine bessere Sakkaden-Pfadplanung für die wiederholte Konfiguration in der visuellen Suche dar. Ähnliche Ergebnisse wurde in einer früheren Studie gezeigt (Zou et al., 2012), in der die Forschungsergebnisse darlegen, dass irrelevante Töne in einer schwierigen visuellen Suchaufgabe zu einer Verlängerung der aktuellen Fixationsdauer führte. Allerdings wurde das Ergebnis der verlängerten Fixationsdauer nicht immer manifestiert. Zum Beispiel beobachteten Tseng and Li (2004) und Zhao et al. (2012) keinen signifikanten Unterschied der Fixationsdauer zwischen den alten und neuen Konfigurationen und Van Asselen et al. (2011) fanden, dass im Vergleich zu der neuen Konfiguration die Fixationsdauer für die alte Konfiguration reduziert wurde. Die unterschiedlichen Ergebnisse über die Fixationsdauer in der Literatur erfordern weitere Forschung.

Zusammenfassend liefert die vorliegende Arbeit mit drei Studien verschiedene neue Beweise, die einige offene Fragen über die Rolle der verschiedenen Kontexte (d.h. globaler, lokaler Kontext und Hintergrundkontext) bei der visueller Suche beantworten. Die Ergebnisse zeigen, dass der räumliche Kontext auf Basis eines rein lokalen Kontexts gelernt werden kann, der innerhalb eines durch die Blickbewegung gesteuerten eingeschränkten zu einem bestimmten Zeitpunkt präsentiert wird. Aber um effektiv den gelernten Kontext abrufen zu können, müssen einige periphere globale visuelle Informationen verfügbar sein (Studie 1). Darüber hinaus haben wir gefunden, dass es von besonders wichtiger Bedeutung ist, die lokalen unveränderten Informationen aufrecht zu halten, um den „Contextual Cueing“-Effekt für den wechselnden Anzeigemodus aufrecht zu erhalten. Der „Contextual Cueing“-Effekt kann nach der Änderung des Anzeigemodus übertragen werden, wenn die Symbole von einem Modus dem anderen Modus gemäß der Methode der „lokalen Unveränderlichkeit" oder „zentralen Unveränderlichkeit" neu zugeordnet wurden (Studie 2). Schließlich fanden wir, dass die Hintergrundinformationen (z.B. ein Hintergrund-Quader) wahrscheinlich in den erlernten Kontext integriert wird, wenn der Hintergrund und Vordergrund stark miteinander assoziiert werden. Jegliche Änderung der Hintergrundinformation wird den Gedächtnisabruf mit Hilfe des „Contextual Cueing“-Effekts verhindern. Aber wenn die Assoziation zwischen Vordergrund und Hintergrund relativ schwach ist, werden die Hintergrundinformationen wahrscheinlich während des kontextuellen Lernens ignoriert (Studie 3).

References

Annac, E., Manginelli, A. A., Pollmann, S., Shi, Z., Müller, H. J., & Geyer, T. (2013). Memory under pressure: Secondary-task effects on contextual cueing of visual search. *Journal Of Vision, 13*(13), 1-15. doi: 10.1167/13.13.6

Assumpção, L., Shi, Z., Zang, X., Müller, H. J., & Geyer, T. (2015). Contextual cueing: implicit memory of tactile context facilitates tactile search. *Attention, Perception, & Psychophysics*.

Baylis, G. C., & Driver, J. (1992). visual parsing and response competition: the effect of grouping factors. *Perception & Psychophysics, 51*(2), 145-162.

Baylis, G. C., & Driver, J. S. (1993). Visual attention and objects: evidence for hierarchical coding of locations. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 451-470.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene Perception: Detecting and Judging Objects Undergong Relational Violations. *Cognitive Psychology, 14*, 143-177.

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Giving a helping hand: effects of joint attention on mental rotation of body parts. *Experimental Brain Research, 211*, 531-545. doi: 10.1007/s00221-011-2625-z

Borst, G., Kievit, R. A., Thompson, W. L., & Kosslyn, S. M. (2011). Mental Rotation is Not Easily Cognitively Penetrable. *Cognitive Psychology, 23*.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: modeling contextual cuing. *Journal of Experimental Psychology:Human Perception and Performance, 33*(4), 798-815. doi: 10.1037/0096-1523.33.4.798

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436. doi: 10.1163/156856897X00357

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. doi: 10.1037/0278-7393.32.4.699

Brockmole, J. R., & Henderson, J. M. (2006a). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *The Quarterly journal of experimental psychology, 59*(7), 1177-1187.

Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*(1), 99-108.

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The Nesting of Search Contexts within Natural Scenes: Evidence from Contextual Cuing. *Journal of Experimental Psychology:Human Perception and Performance, 36*(6), 1046-1048. doi: 10.1037/a0019257

Caputo, G. (1996). The Role of the Background: Texture Segregation and Figure-Ground Segmentation. *Vision Research*, *36*(8), 2815-2826.

Caputo, G., & Casco, C. (1999). A visual evoked potential correlate of global figure-ground segmentation. *Vision Research, 39*, 1597-1610.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review, 6*(2), 204-223.

Chua, K. P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics, 65*, 72-80.

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

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

Chun, M. M., & Jiang, Y. (1999). Top-Down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science*, *10*(4), 360-365. doi: 10.1111/1467-9280.00168

Chun, M. M., & Jiang, Y. (2003). Implicit, Long-Term Spatial Contextual Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(2), 224-234. doi: 10.1037/0278-7393.29.2.224

Chun, M. M., & Nakayama, K. (2000). On the Functional Role of Implicit Visual Memory for the Adaptive Deployment of Attention Across Scenes. *Visual Cognition, 7*(1-3), 65-81. doi: 10.1080/135062800394685

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: news from the front. *Trends In Cognitive Sciences, 2*(10), 1364-1366.

Conci, M., & Müller, H. J. (2012). Contextual learning of multiple target locations in visual search. *Visual Cognition, 20*(7), 746-770. doi: 10.1080/13506285.2012.694376

Conci, M., Müller, H. J., & von Mühlenen, A. (2013). Object-based implicit learning in visual search:Perceptual segmentation constrains contextual cueing. *Journal Of Vision, 13*(3), 1-17. doi: 10.1167/13.3.15

Conci, M., Sun, L., & Müller, H. J. (2011). Contextual remapping in visual search after predictable target-location changes. *Psychological Research, 75,* 279-289. doi: 10.1007/s00426-010-0306-3

Conci, M., & von Muhlenen, A. (2009). Region segmentation and contextual cuing in visual search. *Attention, perception & psychophysics, 71*(7), 1514-1524. doi: 10.3758/APP.71.7.1514

Conci, M., & von Mühlenen, A. (2011). Limitations of perceptual segmentation on contextual cueing in visual search. *Visual Cognition, 19*(2), 203-233.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behav Res Methods Instrum Comput, 34*(4), 613-617.

Daniel, B., & Deubel, H. (2008). Attentional Selection of Multiple Goal Positions Before Rapid Hand Movement Sequences: An Event-related Potential Study. *Journal Of Cognitive Neuroscience*, *21*(1), 18-29.

De Graef, P. (2007). Transsaccadic recognition in scene exploration. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray & R. L. Hill (Eds.), *Eye Movements: A Window on Mind and Brain* (1 ed., pp. 165-191): Elsevier.

Deubel, H., Koch, C., & Bridgeman, B. (2010). Landmarks facilitate visual space constancy across saccades and during fixation. *Vision Research*, *50*(2), 249-259. doi: 10.1016/j.visres.2009.09.020

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold *Psychonomic Bulletin & Review, 4*(1), 3-23.

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001a). Segmentation, attention and phenomenal visual objects. *Cognition, 80,* 61-95.

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001b). Segmentation, attention and phenomenal visual objects. *Cognition, 80,* 61-95.

Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-Contingent Displays: Review and Current Trends. *Cyberpsychology & Behavior, 7*(6), 621-634.

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

Endo, N., & Takeda, Y. (2005). Use of spatial context is restricted by relative position in implicit learning. *Psychonomic Bulletin & Review, 12*(5), 880-885.

Geringswald, F., Baumgartner, F., & Pollmann, S. (2012). Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Frontiers in human neuroscience, 6,* 1-9. doi: 10.3389/fnhum.2012.00134

Geringswald, F., Herbik, A., Hoffmann, M. B., & Pollmann, S. (2013a). Contextual cueing impairment in patients with age-related macular degeneration. *J Vis, 13*(3). doi: 10.1167/13.3.28

Geringswald, F., Herbik, A., Hoffmann, M. B., & Pollmann, S. (2013b). Contextual cueing impairment in patients with age-related macular degeneration. *Journal Of Vision, 13*(3). doi: 10. 1167/13.3.28

Geyer, T., Baumgartner, F., Müller, H. J., & Pollmann, S. (2012). Medial temporal lobe-dependent repetition suppression and enhancement due to implicit vs. explicit processing of individual repeated search displays. *Front. Hum. Neurosci, 6*(272). doi: 10.3389/fnhum.2012.00272

Geyer, T., Shi, Z., & Müller, H. J. (2010). Contextual cueing in multiconjunction visual search is dependent on color- and configuration-based intertrial contingencies. *Journal of Experimental Psychology:Human Perception and Performance, 36*(3), 515-532. doi: 10.1037/a0017448

Geyer, T., Zehetleitner, M., & Müller, H. J. (2010). Contextual cueing of pop-out visual search: when context guides the deployment of attention. *Journal Of Vision, 10*(5), 20, 21-11. doi: 10.1167/10.5.20.

Gibson, E. J. (1969). Principles of perceptual learning and development. Appleton–Century–Crofts 32.

Han, S., Humphreys, G. W., & Chen, L. (1999). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception & Psychophysics, 61*(4), 661-674.

Hartong, D. T., Berson, E. L., & Dryja, T. P. (2006). Retinitis pigmentosa. *Lancet, 368*(9549), 1795-1809. doi: 10.1016/S0140-6736(06)69740-7

Heeger, D. (1997). Signal detection theory. *Dept. Psych., Stanford Univ., Stanford, CA, Teaching Handout*.

Hochstein, S., & Ahissar, M. (2002). View from the Top: Hierarchies and Reverse Hierarchies in the Visual System. *Neuron, 36*(5), 791-804.

Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of Fixation Duration in Visual Search. *Vision Research, 38*(9).

Hooge, I. T. C., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research, 39,* 1567-1575.

Intraub, H. (2002). Visual Scene Perception. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science.* (pp. 524-527). London: Nature Publishing Group.

Ionta, S., & Blanke, O. (2009). Differential influence of hands posture on mental rotation of hands and feet in left and right handers. *Experimental Brain Research, 195*(2), 207-217. doi: 10.1007/s00221-009-1770-0

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194-203. doi: 10.1038/35058500

Itti, L., Koch, C., & Niebur, E. (1998). A Model of Saliency-based Visual Attention for Rapid Scene Analysis. *Ieee Transactions On Pattern Analysis And Machine Intelligence, 20*(11), 1254-1259.

Jeffries, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 54A*(4), 1105-1124.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review, 12*, 100-106.

Jiang, Y., Song, J. H., & Rigas, A. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review, 12*, 524-529.

Jiang, Y., & Swallow, K. M. (2013). Body and head tilt reveals multiple frames of reference for spatial attention. *Journal Of Vision, 13*(13). doi: 10.1167/13.13.9

Jiang, Y., Swallow, K. M., & Capistrano, C. G. (2014). Visual search and location probability learning from variable perspectives. *Journal Of Vision, 13*(6). doi: 10.1167/13.6.13

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing--configuration or individual locations? *Perception and psychophysics, 66,* 454-463.

Jiang, Y., Won, B.-Y., & Swallow, K. M. (2014). First saccadic eye movement reveals persistent attentional guidance by implicit learning. *Journal of Experimental Psychology: Human Perception & Performance*.

Jonikaitis, D., Deubel, H., & de'Sperati, C. (2009). Time gaps in mental imagery introduced by competing saccadic tasks. *Vision Research, 49*(17), 2164-2175.

Kunar, M. A., Flusberg, S., Horowitz, T. S., & Wolfe, J. M. (2007). Does Contextual Cueing Guide the Deployment of Attention. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 816-828.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cueing by global features. *Perception & Psychophysics, 68*(7), 1204-1216. doi: 10.3758/BF03193721

Kunar, M. A., John, R., & Sweetman, H. (2013). A configural dominant account of contextual cueing: Configural cues are stronger than colour cues. *The Quarterly Journal of Experimental Psychology*. doi: 10.1080/17470218.2013.863373

Kunar, M. A., Watson, D. G., Cole, L., & Cox, A. (2013). Negative emotional stimuli reduce contextual cueing but not response times in inefficient search. *The Quarterly Journal of Experimental Psychology*. doi: 10.1080/17470218.2013.815236

Kwon, M., Nandy, A. S., & Tjan, B. S. (2013). Rapid and persistent adaptability of human oculomotor control in response to simulated central vision loss. *Current Biology, 23*(17), 1663-1669. doi: 10.1016/j.cub.2013.06.056

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28*, 1311-1328. doi: 10.1068/p2935

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal Of Experimental Psychology:Applied, 8*(2), 99-117. doi: 10.1037//1076-898X.8.2.99

Loschky, L. C., & McConkie, G. W. G. (2000). *User performance with gaze contingent multiresolutional displays*. Paper presented at the Eye Tracking Research & Applications Symposium 2000

Macmillan, N. A. (2002). Signal detection theory. *Stevens' handbook of experimental psychology*.

Makovski, T., & Jiang, Y. (2010). Contextual cost: when a visual-search target is not where it should be. *Quarterly Journal Of Experimental Psychology, 63*(2), 216-225. doi: 10.1080/17470210903281590

Manginelli, A. A., Baumgartner, F., & Pollmann, S. (2013). Dorsal and ventral working memory-related brain areas support distinct processes in contextual cueing. *Neuroimage, 67*, 363-374.

Manginelli, A. A., Langer, N., Klose, D., & Pollmann, S. (2013). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention, Perception & Psychophysics, 75*(6), 1103-1117. doi: 10.3758/s13414-013-0466-5

Manginelli, A. A., & Pollmann, S. (2009). Misleading contextual cues: how do they affect visual search? *Psychological Research, 73*(2), 212-221. doi: 10.1007/s00426-008-0211-1

Martínez, A., Teder-Sälejärvi, W., Vazquez, M., Molholm, S., Foxe, J., Javitt, D., . . . Hillyard, S. A. (2006). Objects are highlighted by spatial attention. *Cognitive Neuroscience, 18*(2), 298-310.

Mazza, V., Turatto, M., & Umilta, C. (2005). Foreground-background segmentation and attention: A change blindness study. *Psychological Research, 69,* 201-210.

McDonnell, G. P., Mills, M., McCuller, L., & Dodd, M. D. (2014). How does implicit learning of search regularities alter the manner in which you search? *Psychological Research*. doi: 10.1007/s00426-014-0546-8

Melcher, D., & Morrone, C. (2007). Transsaccadic Memory: Building a Stable World from Glance to Glance. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray & R. L. Hill (Eds.), *Eye Movements: A Window on Mind and Brain* (1 ed., pp. 214-233): Elsevier.

Müller, H. J., Heller, D., & Ziegler, J. (1995a). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*(1), 1-17.

Müller, H. J., Heller, D., & Ziegler, J. (1995b). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*(1), 1-17.

Müller, H. J., & Humphreys, G. W. (2003). *A connectionist model of visual search for simple form conjunctions.* Paper presented at the Visual Search 2: Proceedings Of the 2nd International Conferences On Visual Search. CRC Press.

Murphy, K., Torralba, A., Eaton, D., & Freeman, W. (2006). Object Detection and Localization Using Local and Global Features. *Lecture Notes In Computer Science, 4170,* 382-400.

Nakayama, K., Shimojo, S., & Silvermann, G. H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occuluded objects. *Perception, 18*(1), 55-68.

Nakayama, K., & Silvermann, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264-265.

Navon, D. (1977). Forest Before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology, 9*, 353-383.

O'Regan, K., Deubel, H., J., C. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7*, 191-211.

Ogawa, H., & Watanabe, K. (2010). Time to learn: Evidence for two types of attentional guidance in contextual cueing. *Perceptition, 39*, 72-80.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends In Cognitive Sciences, 11*(12), 520-527. doi: 10.1016/j.tics.2007.09.009

Olson, I. R., & Chun, M. M. (2001). Temporal Contextual Cuing of Visual Attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1299-1313. doi: 10.1037//0278-7393.27.5.1299

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9*, 273-302. doi: 10.1080/13506280042000162

Palmer, S. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology, 24*, 436-447.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review, 1*(1), 29-55.

Parkhurst, D., Culurciello, E., & Niebur, E. (2000). *Evaluating Variable Resolution Displays with Visual Search : Task Performance and Eye Movements*. Paper presented at the Proceedings of the 2000 symposium on Eye tracking research & applications, Palm Beach Gardens, Fl, USA.

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

Peterson, M., S. , & Kramer, A., F. . (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63*(7), 1239-1249. doi: 10.3758/BF03194537

Pollmann, S., & Manginelli, A. A. (2009). Early implicit contextual change detection in anterior prefrontal cortex. *Brain Research, 1263*, 87-92. doi: 10.1016/j.brainres.2009.01.039

Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception & Psychophysics*. doi: 10.3758/s13414-013-0446-9

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev, 16*(2), 225-237. doi: 10.3758/PBR.16.2.225

Schlagbauer, B., Müller, H. J., Michael, Z., & Thomas, G. (2012). Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks. *Journal Of Vision, 12*(1).

Schwarzkopf, S. D., & Rees, G. (2010). Interpreting local visual features as a global shape requires awareness. *Proceedings of the Royal society B:Biological Sciences, 278,* 2207-2215. doi: 10.1098/rspb.2010.1909

Schyns, P. G., & Aude, O. (1994). From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science, 5*(4), 195.

Shepard, R. N., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science, New Series, 171*, 701-703.

Shi, Z., Zang, X., Jia, L., Geyer, T., & Müller, H. J. (2013). Transfer of contextual cueing in full-icon display remapping. *Journal Of Vision, 13*(3), 1-10. doi: 10.1167/13.3.2

Shomstein, S., & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. *Psychonomic Bulletin & Review, 11,* 247-253.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit test. *Memory & Cognition, 36*(2), 403-415. doi: 10.3758/MC.36.2.403

Song, J. H., & Jiang, Y. (2005). Connecting the past with the present: how do humans match an incoming visual display with visual memory? *Journal Of Vision, 5*, 322-330. doi: 10.1167/5.4.4

Torralba, A., Oliva, A., & Castelhano, M. S. (2006). Contextual Guidance of Eye Movements and Attention in Real-World Scenes: The Role of Global Features in Object Search. *Psychological Review, 113*(4), 766-786.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, & Image Processing, 31*, 156-177.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review, 95*, 15-48.

Treisman, A. N., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology, 12*, 97-136.

Tseng, Y.-C., & Li, C.-S. R. (2004). Oculomotor correlates of context-guided learning in visual search. *Perception & Psychophysics, 66*(8), 1363-1378. doi: 10.3758/s13414-012-0400-2

Tseng, Y.-C., & Lleras, A. (2013). Rewarding context accelerates implicit guidance in visual search. *Attention, perception & psychophysics, 75*(2), 287-298. doi: 10.3758/s13414-012-0400-2

Van Asselen, M., & Castelo-Branco, M. (2009). The role of peripheral vision in implicit contextual cuing. *Attention, Perception, & Psychophysics, 71*(1), 76-81.

Van Asselen, M., Sampaio, J., Pina, A., & Castelo-Branco, M. (2011). Object based implicit contextual learning: a study of eye movements. *Attention, perception & psychophysics, 73*(2), 297-302. doi: 10.3758/s13414-010-0047-9

Westerberg, C. E., Miller, B. B., Reber, P. J., Cohen, N. J., & Paller, K. A. (2011). Neural correlates of contextual cueing are modulated by explicit learning. *Neuropsychologia, 49*(12), 3439-3447.

Wolfe, J. M. (1994a). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202-238. doi: doi:10.3758/BF03200774

Wolfe, J. M. (1994b). Guided Search 2.0 A revised model of visual search. *Psychon Bull Rev, 1*(2), 202-238. doi: 10.3758/BF03200774

Wolfe, J. M. (2003a). Moving towards solutions to some enduring controversies in visual search. *Trends In Cognitive Sciences, 7*(2).

Wolfe, J. M. (2003b). Moving towards solutions to some enduring controversies in visual search. *Trends In Cognitive Sciences, 7*(2), 70-76.

Zang, X., Jia, L., Müller, H. J., & Shi, Z. (2013). Invariant spatial context is learned but not retrieved in gaze-contingent limited-viewing search. *Journal of Experimental Psychology: Learning, Memory, and Cognition (Submitted).*

Zellin, M. (2012). *Implicit Adaptation to Change in Memory-Guided Visual Search.* (Dr. Phil.), Ludwig-Maximilians-Universität, München.

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2011). Two (or three) is one too many: testing the flexibility of contextual cueing with multiple target locations. *Attention, Perception & Psychophysics, 73*(7), 2065-2076. doi: 10.3758/s13414-011-0175-x

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). doi: 10.1371/journal.pone.0059466

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).

Zellin, M., Von Mühlenen, A., Müller, H. J., & Conci, M. (2014). Long-term adaptation to change in implicit contextual learning. *Psychonomic Bulletin & Review*, 1-7. doi: 10.3758/s13423-013-0568-z

Zhao, G., Liu, Q., Jiao, J., Zhou, P., Li, H., & Sun, H.-j. (2012). Dual-state modulation of the contextual cueing effect: Evidence from eye movement recordings. *Journal Of Vision, 12*(6), 11:11-13. doi: 10.1167/12.6.11

Zou, H., Müller, H. J., & Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision., 12*(5), 2: 1-18. doi: 10.1167/12.5.2

Acknowledgements

The studies reported in this thesis were written at the Ludwig-Maximilians University in München, and I have been financially supported by the German Research Council (DFG) research project grant SH166 and China Scholarship Council (CSC).

The decision to start my doctoral study in Psychology three years ago is one of the most important and wisest choices that I have ever made. Over the last three year, I have been working on my favourite research topic, enjoying a beautiful life. And, most importantly, I have met many great friends and colleagues who helped me a lot in both my academic research and personal life. In particular, I would like to thank Prof. Dr. Hermann J. Müller who offered me a position in the psychology department, gave me fully support throughout my research work, offered inspiring meeting in Herrsching/Neusäß, and being able to discuss and answer all my questions related to the research.

The thesis would not have been possible without help, support and patience of my supervisor PD. Dr. Zhuanghua Shi, who led me into the area of psychology. He has been supportive and has given me the freedom to pursue various research projects, and has helped me on every step of my whole doctoral research. Furthermore, I would like to thank my second supervisor, PD. Dr. Thomas Geyer, who has excellent personality and scientific expertise. He is always gentle and willing to answer my scientific questions. I am so grateful that Thomas encouraged me almost every time we met. His inspiration helped me a lot and made me never lack of passion in my doctoral work.

I would like to thank Prof. Dr. Dongheui Lee from Technische Universität München, who is kindly willing to be one of my examiners. Many thanks also go to my former

supervisors, Dr. Drazen Brscic, Prof. JianPing Wang, and Prof. Dr. Sandra Hirche for their support and kindness of my study and research. I also thank to my colleagues and friends, Leonado Assumpcao, Dr. Lina Jia, Julia Alice Ewerdwalbesloh, Ondrej Havlicek, Kelsey Carlson, Lena Schröder, Omiros Kourakos, Vasiliki Koropouli, Xiuna Zhua, and many other friends that are not listed here for because they are able to give me a hand in time when I encounter problems in my research and life. Especially thanks go to Mengtao Gu, Leonado, Julia, Lena and Luciana Chong, who provide me supports in improving English and German.

Last but not least, I would like to thank my dear parents, lovely brothers and my boyfriend, who continuously encourage and support me during the past years. Your love and support means the world to me!

Xuelian Zang

May, 2014

Curriculum Vitae

Xuelian Zang

Born on 28th April 1986 in Anhui Province, China

Education

2011-2014	Doctoral student in Psychology, Ludwig-Maximilians-Universität, München
2007-2009	Master in Measurement Technology and Automation Equipment, China Agricultural University
2003-2007	Bachelor in Automation, China Agricultural University

Professional Experience

2011-2014	Research Assistant, Department of Psychology, Ludwig-Maximilians-Universität, München
2009-2011	Research Assistant, Department of Electrical Engineering and Information Technology, Technische Universität München

Scholarships

2013	European Conference on Visual Perception Travel Award;
2009-2013	Doctoral scholarship from Chinese Scholarship Council (CSC).
2005-2007	National scholarship sponsored by the Ministry of Education, China.
2004-2007	Scholarships for outstanding students, China Agricultural University.

Selected Publications

Assumpcao, L., Shi, Z**., Zang, X**., Müller, H. J., Geyer, T. (2014). Contextual cueing: implicit memory of tactile context facilitates tactile search. (Attention, Perception, & Psychophysics, in press).

Zang, X., Jia, L., Müller, H. J., Shi, Z. (2014.10), Invariant spatial context is learned but not retrieved in gaze-contingent limited-viewing search. Journal of Experimental Psychology: Learning, Memory, and Cognition (in press).

Shi, Z., **Zang, X**., Jia, L., Geyer, T., Müller, H. J. (2013). Transfer of contextual cueing in full-icon display remapping. Journal of Vision, 13(3):2, 1–10, doi:10.1167/13.3.2.

Jia, L., Shi, Z., **Zang, X**., Müller, H. J. (2013), Concurrent emotional pictures modulate spatial-separated audio-tactile temporal order judgments, Brain Research, 1537:156-163.

Zang X, Bršcic D, Wang W. Kühnlenz K. (2012), Human behavior inspired temporal attention control system for multiple object tracking. Journal of University of Science and Technology of China, 2012, 42(9): 714-722.

Zang X., Bršcic D., Wang W., Kühnlenz K. (2012), Temporal attention control system for multiple objects tracking, The 31st Chinese Control Conference (CCC'2012), p. 4836-4841.

Appendix

Journal of Vision (2013) 13(3):2, 1–10 http://www.journalofvision.org/content/13/3/2

Transfer of contextual cueing in full-icon display remapping

Zhuanghua Shi
Experimental Psychology, Department of Psychology,
Ludwig-Maximilians-Universität München,
Munich, Germany

Xuelian Zang
Experimental Psychology, Department of Psychology,
Ludwig-Maximilians-Universität München,
Munich, Germany

Lina Jia
Experimental Psychology, Department of Psychology,
Ludwig-Maximilians-Universität München,
Munich, Germany

Thomas Geyer
Experimental Psychology, Department of Psychology,
Ludwig-Maximilians-Universität München,
Munich, Germany

Hermann J. Müller
Experimental Psychology, Department of Psychology,
Ludwig-Maximilians-Universität München,
Munich, Germany
School of Psychological Science, Birkbeck College
(University of London), London, UK

Invariant spatial context can expedite visual search, an effect that is known as contextual cueing (e.g., Chun & Jiang, 1998). However, disrupting learned display configurations abolishes the effect. In current touch-based mobile devices, such as the iPad, icons are shuffled and remapped when the display mode is changed. However, such remapping also disrupts the spatial relationships between icons. This may hamper usability. In the present study, we examined the transfer of contextual cueing in four different methods of display remapping: position-order invariant, global rotation, local invariant, and central invariant. We used full-icon landscape mode for training and both landscape and portrait modes for testing, to check whether the cueing transfers to portrait mode. The results showed transfer of contextual cueing but only with the local invariant and the central invariant remapping methods. We take the results to mean that the predictability of target locations is a crucial factor for the transfer of contextual cueing and thus icon remapping design for mobile devices.

Introduction

Invariant visual context provides an important spatial cue for the guidance of visual search and focal-attentional selection. Repeated exposure to the same arrangements of search displays facilitates reaction time (RT) performance, an effect that has been referred to as contextual cueing (Chun, 2000; Chun & Jiang, 1998; Chun & Nakayama, 2000). In their seminal paper, Chun and Jiang (1998) had their observers search for a target letter "T" embedded in a set of distractor letters "L". Unbeknown to participants, half of the presented displays contained identical configurations of target and distractor items (i.e., old displays), whereas the other half contained novel configurations (i.e., new displays). The main result was that of faster RTs to old relative to new displays (i.e., contextual cueing), an effect that developed after a short period of training. Interestingly, when observers were queried about repeated displays at the end of the search task in an "old-new" recognition test, their performance was only at chance level. From these findings, Chun and

Citation: Shi, Z., Zang, X., Jia, L., Geyer, T., & Müller, H. J. (2013). Transfer of contextual cueing in full-icon display remapping. *Journal of Vision*, *13*(3):2, 1–10, http://www.journalofvision.org/content/13/3/2, doi:10.1167/13.3.2.

doi: 10.1167/13.3.2 Received October 9, 2012; published February 25, 2013 ISSN 1534-7362

Jiang (1998) concluded that (a) contextual cueing guides focal attention more rapidly to the target location (but see Kunar, Flusberg, Horowitz, & Wolfe [2007], for evidence that contextual cueing might also aid postperceptual processes) and (b) the cueing effect derives from an implicit memory for the items' spatial arrangement. Since then, the cueing effect has been elaborated in a number of further studies (Chun, 2000; Chun & Jiang, 1998; Chun & Nakayama, 2000; Conci, Sun, & Müller, 2011; Conci & von Mühlenen, 2009, 2011; Geyer, Shi, & Müller, 2010; Jiang & Wagner, 2004; Kunar, Flusberg, & Wolfe, 2006). Jiang and Wagner (2004; see also Brady & Chun, 2007, or Olson & Chun, 2002) showed that contextual cueing is supported by two distinct spatial memory systems for individual item locations (i.e., local learning) and, respectively, the entire configuration formed by the distractors (i.e., global learning). Further, Kunar et al. (2006) showed that nonspatial attributes, too, such as background color, can facilitate RT performance. Contextual learning is also influenced by selective attention: Only the arrangement of some items, in particular, those sharing the target color, are learned over the course of an experiment (e.g., Geyer et al., 2010; Jiang & Leung, 2005).

However, the degree to which contextual cueing can adapt to changes in learned displays remains subject to debate. For example, Jiang and Wagner (2004) reported that contextual cueing was still reliable even when learned displays were shifted along the horizontal display axis, the vertical display axis, or presented in a different size (compressed or expanded). Other studies (Brady & Chun, 2007; Olson & Chun, 2002) showed that contextual cueing survived changes of approximately 50% up to 75% of the display items; that is, cueing was reliable even when only one half or one quadrant of the display was repeated across trials. On the other hand, Olson and Chun (2002) reported that the cueing effect was abolished when new distractors were presented in between the target and the old distractors, with the target being presented, for example, in the left half and the old distractors in the right half of the display. Several other studies confirmed that contextual cueing diminished when the target was repositioned in repeated displays and thus became unpredictable (Chun & Jiang, 1998; Manginelli & Pollmann, 2009; Olson & Chun, 2002; Wolfe, Klempen, & Dahlen, 2000). In contrast, the contextual cueing effect remained effective with predictable target location changes (Conci & Müller, 2012; Conci et al., 2011). Makovski and Jiang (2010) suggested that predictability based on invariant context is a key factor for contextual cueing, based on their finding that the cueing effect decreased as the target appeared further away from its learned location; in fact, there were even RT costs when the target swapped its location with a previous distractor. Similar findings have been reported in three-dimensional (3D) scene search (Chua & Chun, 2003), in which contextual cueing decreased with increasing angular difference between viewpoints in the training versus the test displays (the experiment was divided into a training and test phase, with the latter containing modified displays).

Although most of the work on contextual cueing was conducted using consistent (i.e., spatially invariant) search displays with a fixed number of items (e.g., one target and 11 distractors presented at a total of 48 locations within an invisible 6 × 8 matrix), none of these studies has examined the influence of changes of the display orientation on the cueing effect. Although changing display mode (and accordingly remapping of the items) occurs rarely with standard (i.e., laboratory) displays, switching display mode is a normal routine in current touch-based mobile devices, such as the iPad. Interestingly, with these devices, there is only one type of item—or icon—remapping method available: The positions of icons in one display (e.g., landscape mode) are remapped to the other display (portrait mode) by keeping the positional order (left to right and up to down) constant across all icons (see Figure 1a, b). Although this remapping method preserves the positional order and 80% of the horizontal intericon relationships (in a 4 × 6 icon matrix, as shown in Figure 1a, b), it destroys almost all local icon relationships, in particular, when the display is arranged as a rectangle (as with almost all mobile devices). However, based on the contextual cueing studies reviewed above, it is possible that contextual cueing is reduced, if not entirely abolished, when display orientation changes from landscape to portrait mode and icons are remapped in the standard position-order manner. Given this, one intriguing question arises, namely, are there any other improved methods for icon remapping, such that the remapping could enhance users' performance in everyday situations of display mode changes? This question was addressed in the current study by using the contextual cueing effect as a tool to evaluate the effectiveness of various display-remapping techniques; that is, preserved contextual cueing from one to the other display mode was taken as an indicator for the value of a given remapping method.

Besides the position-order remapping method, several other (simple) remapping methods are possible. For example, one of the most natural ways is to rotate the entire display by 90° in the clockwise direction (individual icons are rotated 90° in the counterclockwise direction to keep their appearance constant; see Figure 1c). Such a global rotation is similar to the rotation of an object in our physical world (e.g., imagine you rotate a key cabinet with many keys). Alternatively, and motivated by the above mentioned studies on contextual cueing (e.g., Brady & Chun, 2007), one could also try to preserve local associations

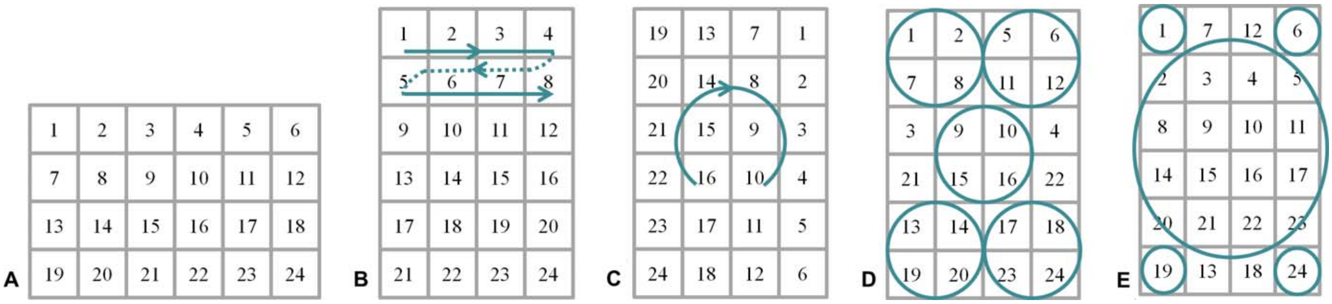

Figure 1. Schematic illustration of display layouts and remapping methods. (A) Display layout in landscape mode; each number denotes an individual icon. (B) Portrait display layout obtained by the position-order invariant remapping method; the arrow indicates the icon remapping sequential order from the landscape to portrait mode. (C) Portrait display layout obtained by the global rotation remapping method; the arrow indicates the rotation direction from the landscape to portrait mode. (D) Portrait display layout obtained by the local invariant remapping method; circled regions remain the same between the landscape and portrait mode. (E) Portrait display layout obtained by the central invariant remapping method; circled regions are invariant.

within the entire configuration as completely as possible. There are two ways to maximize such local invariants. One is to subdivide displays into several local regions and preserve the placement of these local regions in the entire configuration after icon remapping (Figure 1d). Another method is to keep the display center constant in remapped displays (Figure 1e).

To investigate how these various display-remapping methods influence memory performance, we examined contextual cueing effects in four separate experiments. Each experiment examined one display remapping method. To simulate touch-based icon displays and observers' active touch action, we used real desktop icons as search items and presented them on a touch monitor in the four experiments.

Methods

Participants

A total of 40 observers took part in the experiments (10 in each experiment, mean ages: 27.9, 26.2, 25.5, and 27.3 years and number of females: 7, 6, 6, and 5 for Experiments 1–4, respectively). All had normal or corrected-to-normal visual acuity (including color vision). They gave written consent prior to the experiment and were paid at a rate of €8/hour for taking part. Participants were naive as to the intention of the study.

Apparatus and stimuli

The experiments were conducted in a dimly lit cabin (ambient light: 4.36 cd/m^2). Visual stimuli were presented on a 23-inch multitouch LCD monitor (HP2310ti) with spatial resolution set to 1920 × 1080 pixels. To make touch pointing comfortable for the participants, the screen panel was placed on the table tilted by 45°. The viewing distance was approximately 40 cm, with participants' head position fixed by a chin rest. Twenty-four typical computer icons (randomly selected from 48 candidate icons[1] for each observer) were presented within an invisible 6 × 4 horizontal grid (subtending 24° × 16° of visual angle) or a 4 × 6 vertical grid (subtending 24° × 16°). The target was the icon with a top overlay of a compound letter "T" (subtending 1.6° × 1.6°; luminance 35.67cd/m^2; see Figure 2). Such a compound target letter was used for two reasons: first, to avoid interference between the target and some other (distractors) letters, and second, to make the compound letter and the icon comparable in terms of their luminance level. The background of the search displays was set to gray (16.56 cd/m^2). To enhance the global spatial "Gestalt" (i.e., perception of the display as landscape or, respectively, portrait mode), we added one array of six upright white triangles (130.5 cd/m^2) with a gray background (19.62 cd/m^2) below the landscape mode (Figure 2a) or to the left side of the portrait mode (Figure 2b). The triangle array was meant to serve as a global landmark in the experiments, indicating display mode changes. The experimental program was developed with and controlled by Matlab (Mathworks Inc., Natick, MA), in addition to the Psychtoolbox (Brainard, 1997; Pelli, 1997). Response times were recorded via the touch screen. To determine the onset of a response, an additional input button (connected to a NI PXI system) was placed in between the touch screen and the participants, which was used for initiating the task and pointing movement.

Design and procedure

A three-factorial within-subject design was used with display mode (landscape, portrait), context (old, new),

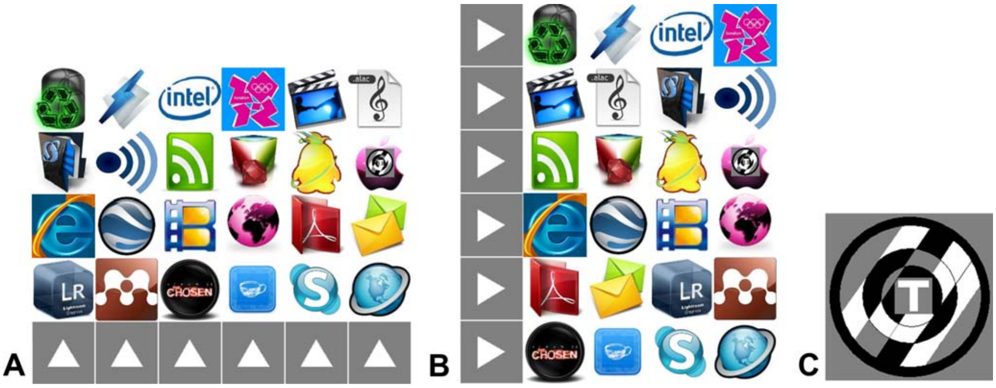

Figure 2. Example displays in the experiments. (A) Example of a landscape display. In this example, the "Apple" icon (second row, right-most column) is the search target. (B) Example of a portrait display. In this case, icons are remapped from the landscape mode by keeping the position order constant in the left-to-right and up-to-down manner (Experiment 1). (C) The top overlay for the target icon (a compound letter "T").

and experimental epoch (1–9) as independent variables. From the 24 possible target locations, we randomly selected 12 target positions for old and the other 12 positions for new displays. In this way, the target appeared equally likely at any of the 24 possible locations. To have enough difference between old and new configurations and to control the similarity of icon identities, we selected 24 icons from 48 typical icon candidates and assigned to random locations. Each of the new target locations was paired with newly generated distractor icons for every new-display trial, whereas each of the old target locations was paired with randomly selected distractor icons at the beginning of each experiment and served as old landscape displays. These old landscape displays were also used to define the remapped old portrait displays. Remapping was one as follows:

a. Experiment 1 (*position-order invariant*). The positional order (left to right and top to bottom) of the icons in the portrait mode was the same as that in the landscape mode (Figure 1b). This method is used in most of the present mobile devices for the rearrangement of icons.
b. Experiment 2 (*global-rotation*). The landscape display was, as a whole, rotated by 90° clockwise into the portrait mode, while preserving the (upright) orientation of the individual icons. With this global rotation, the global and local relationships of the icons are rotated by 90° across display changes (Figure 1c).
c. Experiment 3 (*local invariant*). To preserve the local (and global) spatial configuration as much as possible, in Experiment 3, the display was divided into four peripheral and one central region, each consisting of four icons (see circled regions in Figure 1d). The positioning of these four "corners" and the central region were kept constant across display mode changes. Only four remaining items (i.e., Icons 3, 4, 21, and 22 in Figure 1d) changed their relative positions. Similar to the global rotation, with the local-invariant transformation, the local relationships between all icons are preserved across display changes.
d. Experiment 4 (*central-invariant*). Instead of dividing the display into multiple regions, in Experiment 4, we preserved the central display region as much as possible (i.e., preserving the central maximum square region). As shown in Figure 1e, icons in the central 4 × 4 matrix were positioned at identical locations across display mode changes. In addition, the four outermost (corner) icons were also unchanged. Only the remaining four icons (7, 12, 13, and 18 in Figure 1e) changed their positions.

Each experiment comprised three consecutive sessions: learning, test, and recognition. In the learning session, there were five epochs of three blocks, with each block consisting of 24 search trials. To keep the experiment as short as possible, the learning session contained only 12 old-landscape displays to foster learning effect (each of the old display repeated twice per block). The transfer session had four epochs, with each epoch consisting of 24 trials (i.e., one block only). In half of these trials, an old display was presented and new displays in the other half. New displays were randomly generated at the beginning of each trial. The order of display modes in the transfer epochs was fixed: landscape (L), portrait (P), portrait (P), and landscape (L). The first transfer epoch with the landscape mode (i.e., nontransformed) was intended to test for a standard contextual cueing effect. The last transfer epoch was intended for examining whether contextual cueing is still manifested by two intervening epochs containing different display modes. To avoid confounding by repetition effects, we randomly presented trials in such a way that the same old display was never repeated within three consecutive trials.

In the learning and test sessions, each trial started with a cross-fixation presented in the center of the display. Participants had to press the input button (also serving as the initial hand position) to trigger the presentation of the search display. Participants were instructed to detect the target and touch its location with their index finger as rapidly and as accurately as possible. A blank screen was presented after the localization response, or 4.5 s when no response was made. When participants made an erroneous response, an additional feedback display containing a stop warning sign was presented for 1.0 s. After 1.0 to 1.2 s of intertrial interval, the next trial started.

In the recognition session, participants were asked if they had realized any display repetitions during the learning and transfer sessions and, if so, when they had first noticed the repeated displays (note that a similar protocol was used by Chun & Jiang, 1998). Following this, they had to judge a total of 24 displays, including 12 new displays (six landscape and six portrait displays) and 12 old displays (six landscape and six portrait displays), in an "old-new" recognition test. In this test, the chance rate for recognizing a repeated display was 50%.

Prior to the experiment, participants practiced the experimental task in one training block of 24 trials (data not recorded). The search displays used in the practice trials were not shown later in the experiment. Participants were allowed to take a break in between successive blocks of the experiment. The break between the learning and transfer sessions was similar to other between-block breaks.

Results

Accuracy performance

Error rates were overall small (<1%) and were comparable across all experiments. For further RT data analyses, we excluded trials with erroneous responses and RTs outside the range of 200 to 3000 ms. Such outliers were also low in general (<3%).

Perceptual learning

The mean RTs for the learning sessions are shown in Figure 3 (Epochs 1–5). For each experiment, the mean RTs were examined by repeated-measures analysis of variance (ANOVA) with the single-factor epoch. The main effect was significant for all four experiments (all p's < 0.05); further Bonferroni tests revealed a significant perceptual learning effect, defined as the difference in RTs between Epoch 5 (i.e., the end of the training session) and Epoch 1 (i.e., the beginning of the training session; Table 1). In addition, to examine interference by the introduction of new (both landscape and portrait) displays in the transfer session, RTs for the old displays in the first epoch of the transfer session (Epoch 6) were compared with RTs in the last epoch of the learning session. Although RTs were numerically longer in Epoch 6 compared with Epoch 5, the slowing was significant only for Experiment 2 (Table 2). This suggests that introducing novel displays had only some moderate influence on the search task response.

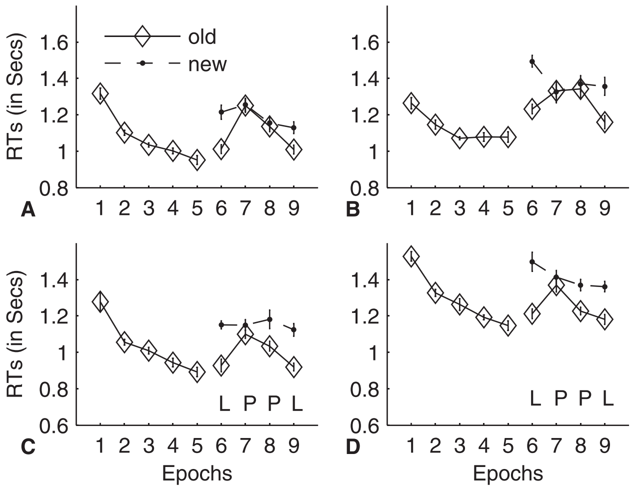

Figure 3. Mean correct response times (RTs) as a function of epoch for the learning (Epochs 1–5) and transfer (Epochs 6–9) sessions; for the latter, mean RTs are shown separately for old displays (denoted by diamonds and solid lines) and new displays (denoted by dots and dashed lines). (A) Experiment 1, position-order invariant remapping. (B) Experiment 2, global rotation remapping. (C) Experiment 3, local invariant remapping. (D) Experiment 4, central invariant remapping.

Transfer of contextual cueing effect

The mean RTs, separately for old and new contexts, as a function of epoch for the test phase are presented in Figure 3 (Epochs 6–9). To examine the contextual cueing effect, mean RTs were subjected to a repeated-measures ANOVA with epoch (6–9) and context (old vs. new) as factors, separately for each experiment. The results are summarized in Table 2. The RTs were significantly faster for old displays compared with new displays in all four experiments, indicating robust contextual cueing benefits. The main effect of epoch was also significant for Experiments 1, 3, and 4, indicating that some perceptual learning also occurred in the transfer session. Finally, the context × epoch interaction was significant for all experiments, reflecting differential cueing effects in the different epochs. Post hoc tests revealed significant contextual cueing to

	Perceptual learning		Interference associated with the presentation of new displays	
Experiment	Facilitation (ms)	ANOVA	Cost (ms)	ANOVA
1	364	$p < 0.01$	60	$p = 0.08$
2	185	$p < 0.05$	154	$p < 0.05$
3	385	$p < 0.01$	36	$p = 0.21$
4	380	$p < 0.01$	98	$p = 0.11$

Table 1. Mean learning effect in the training sessions and interference by the addition of new displays in the transfer session, for each experiment.

be significant for all landscape displays (Epochs 6 and 9). By contrast, for portrait displays (in Epoch 8), significant contextual cueing was evident only in Experiments 3 and 4 (see Table 2). Note that each epoch in the transfer session contained only 24 trials, suggesting that the contextual cueing effect could be quickly transferred with the local invariant and central invariant remapping methods when the display mode was changed.

To examine whether contextual cueing effects were comparable among the different experiments, a repeated-measure ANOVA was conducted on the cueing effect in the first transfer Epoch 6 (with landscape mode), with the single-factor experiment. The effect of experiment was nonsignificant, $F(3, 27) = 0.55$, $p = 0.65$, suggesting that the contextual cueing effects were comparable among experiments. Thus, any differences in the subsequent transfer epochs are likely attributable to the particular method of display (icon) remapping.

Recognition test

Based on participants' postexperimental reports, we determined the percentages of participants who noticed display repetitions during the search task and who attempted to explicitly learn the displays; the times (in terms of the number of blocks performed) at which these participants first noticed the repetitions were also calculated. We then further calculated participants' mean hit and false alarm rates as well as their discrimination sensitivities (d′) for landscape (L) and portrait (P) displays. The results are summarized in Table 3.

In all experiments, participants exhibited high proportions of recognized displays. The recognition sensitivities (d's) were significantly larger than zero for both landscape and portrait displays ($p < 0.05$), except for one marginally significant effect for the landscape display in Experiment 3 ($p = 0.066$), which was mainly due to one observer who showed an extreme negative d_H' score (−1.40). When excluding this participant, d_H' was also significant: $p < 0.05$. Taken together, the significantly positive d′ scores suggest that after learning, participants recognized not only the old landscape displays but also the remapped portrait displays in all four experiments. Moreover, there was no significant difference in recognition sensitivity between the landscape and portrait displays, at least for the first three experiments (see the last column in Table 1), indicating that remapping did not hamper explicit recognition. Although recognition accuracy was lower for portrait than for landscape displays in Experiment 4, the effect was mainly due to the very high recognition sensitivity in the landscape mode (Table 1). Nevertheless, even in Experiment 4, the sensitivity for the portrait displays was still significantly greater than zero, supporting the idea that the transformed old portrait displays can be recognized explicitly. The lack of differential recognition sensitivities between landscape and portrait displays in Experiments 1 and 2 is in contrast to the differential contextual cueing effects with landscape versus portrait displays. This suggests that recognition and visual search may involve different

	Contextual cueing effect (ms)					ANOVA test with *F* value		
Experiment	Average (Epoch 6–9)	Epoch 6	Epoch 7	Epoch 8	Epoch 9	Context (old/new)	Epoch (6–9)	Interaction
1	86	202***	5	18	120*	9.25*	9.23**	4.49*
2	120	262***	7	28	197**	12.48**	1.40	8.80***
3	156	223**	50	147*	205**	21.49***	5.16**	3.07*
4	163	286**	43	142*	190**	22.42***	5.63**	3.91*

Table 2. Contextual cueing effects in the transfer session. The reported significance values are as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Experiment	Noticed repetition	Explicit learning	When (blocks)	Hit rates	False alarms	d'_L	d'_P	$d'_L = d'_P$
1	90%	60%	6.89	76.3%	38.3%	2.17**	1.17*	$p = 0.17$
2	80%	30%	7.25	72.9%	43.3%	1.33**	1.17*	$p = 0.68$
3	90%	20%	5.85	65.1%	28.3%	1.56	2.02*	$p = 0.51$
4	80%	60%	4.75	80%	28.3%	3.42**	1.52*	$p < 0.05$

Table 3. Results of recognition test. The reported significance values are as follows: *$p < 0.05$; **$p < 0.01$.

memory processes, with the former recruiting more complex information-matching processes that do not benefit the search processes.

Discussion

The present study examined the transfer of learned contextual cues in full-icon display remapping. The main goal was to investigate whether contextual cueing continues to facilitate icon localization (RT) performance following display mode changes. We compared four different types of icon remapping: position-order invariant, global-rotation, local-invariant, and central-invariant remapping. In all experiments, robust learning effects were found in the training session for the landscape displays. The RTs were faster at the end relative to the beginning of the training session. This practice effect is likely attributable to general learning of the localization task (Schneider & Shiffrin, 1977).

In the test session, in which new displays were introduced (in addition to the old displays), we established a contextual cueing effects in all experiments, at least when the display mode was kept the same. This suggests that icon identities and spatial configurations among icons could serve as context cues to facilitate the localization task. Note that the facilitation effect might also be partially due to position-based learning, given that only old displays were used in the training session. However, the transfer effects found in the portrait displays (Experiments 3 and 4) cannot be fully explained by position-based learning, because the positions were changed in the portrait displays and positional repetitions were equated between the old and new displays. Interestingly, contextual cueing was evident for landscape displays even after the insertion of two epochs of portrait displays. This may be taken to indicate that the cueing effect is relatively robust against interference within the same set of old configurations, consistent with previous studies (Chun & Jiang, 1998, 2003; Conci et al., 2011; Conci & Müller, 2012; Jiang, Song, & Rigas, 2005; Song & Jiang, 2005; Zellin, Conci, von Mühlenen, & Müller, 2011). However, contextual cues acquired with landscape displays were transferred to portrait displays only under certain remapping conditions (those of Experiments 3 and 4), suggesting that contextual cueing is relatively inflexible and that transfer is confined to specific remapping situations.

The differential pattern of effects revealed among the four experiments raises the question as to the factors that modulate the transfer of learned displays. The position-order invariant method maintained icons in their same left-to-right and up-to-down manner. Although 80% of the horizontal relationships are preserved with this transformation, it destroys almost all vertical relationships. It also changes the absolute positions of the icons dramatically; for instance, Position 5 is shifted from the left side in the landscape display (Figure 1a) to the right side in the portrait display (Figure 1b). As a result, the target location might become unpredictable in remapped displays, abolishing the contextual cueing effect (Conci et al., 2011; Manginelli & Pollmann, 2009). Note that in the current terms, predictability refers to both the target's absolute location on the screen as well as its placement within the entire configuration (given that we did not vary the target's absolute and relative location independently).

When comparing the position-order invariant to the global-rotation method, the latter maintains all local icon neighborhood relationships, but the overall Gestalt is rotated by 90° from the landscape to portrait mode. With this type of remapping, repeated displays failed to facilitate RT performance in portrait displays. Possibly, the contextual associations learned in the landscape displays were quite instance specific and too weak for the global-rotation remapping. As shown in mental rotation studies (Böckler, Knoblich, & Sebanz, 2011; Borst, Kievit, Thompson, & Kosslyn, 2011; Ionta & Blanke, 2009; Shepard & Metzler, 1971; Shomstein & Yantis, 2004), RTs increase linearly with increasing angular disparity when participants were asked to decide whether two presented objects are the same. Those paired objects were normally rotated objects or mirrored objects, and participants had to carry out mental rotation (rotating one object into the other) to solve the task. Applied to the current Experiment 2, although the global rotation maintains the local icons' neighborhood relationships, the mapping of a new portrait onto an old landscape display may likewise be a demanding (i.e., time-consuming) process, which diminishes any performance gains brought about by contextual cueing. In a previous study using 3D visual search, Chua and Chun (2003) also showed that contextual cueing decreased with increasing angular difference between viewpoints of training and test displays. Thus, demanding mental rotation might be the main reason why we failed to find any transfer of contextual cueing from the landscape to the portrait in Experiment 2. It should be noted, however, that in our setup, the experimental program presented the rotated portrait display automatically. That is, participants passively viewed the search displays, rather than carrying out the rotation actively. It would be interesting to examine the transfer of contextual cueing when participants rotate the displays themselves (i.e., actively).

In contrast to Experiments 1 and 2, we found significant transfer of contextual cueing in Experiments 3 and 4, in which the portrait display was remapped from the landscape display using the local-invariant (Experiment 3) or central-invariant methods (Experiment 4). Both experiments disclosed numerical contextual cueing benefits already in the first epoch with portrait displays (50.8 and 44 ms for Experiments 3 and 4, respectively), although these effects were not significant. No contextual cueing in the first portrait epoch is likely due to the orientation change of the whole display. Mapping old landscape to portrait displays may engage additional mental processes, diminishing the contextual cueing effect. In addition, interobserver variability was large because both the old and new displays were presented only once in this epoch. Interestingly, transfer of contextual cueing was highly reliable for both remapping methods (147.4 and 142.0 ms for Experiments 3 and 4, respectively) in the second epoch. The local-invariant remapping method keeps five of seven local regions unchanged, and the global topological relationship of these five local regions also remains the same. This means that local regions appear at the very same positions (quadrants) in the entire configuration after the remapping. Likewise, the central-invariant remapping method maintains the absolute icon positions of the four outermost corners and the central region (83% in total). In both cases, after the remapping, the target position is much more predictable compared with both the position-order invariant and the global rotation methods. In contrast to previous investigations of contextual cueing, suggesting that only three to four repeated items (among some eight novel items) can produce the effect (Song & Jiang, 2005), the results of the present Experiments 1 and 2 suggest that merely preserving some local invariant information does not guarantee transfer of contextual cueing. Instead, the three to four items would have to appear at the very same positions within the global configuration to observe contextual cueing (Experiments 3 and 4; see also Brady & Chun, 2007, for a related proposal, albeit using different approach).

The recognition tests showed that in all experiments, participants were well able to discern repeated from nonrepeated displays. This contrasts with standard contextual cueing studies in which recognition accuracy was typically at chance level (Chun & Jiang, 1998). Explicit memory effects may be due to the heterogeneous and, importantly, realistic icons used as distractors in our experiments (see also Brockmole, Castelhano, & Henderson, 2006). Interestingly, in all the experiments of the present study, recognition accuracy was larger than chance for all landscape and, importantly, remapped portrait displays. In contrast, transfer of contextual cueing was observed only in Experiments 3 and 4. This argues that merely recognizing a repeated display as an old one does not necessarily mean that this also facilitates RT performance. Of interest in this regard, it has been reported that explicit learning of repeated displays engages neural processes that are distinct from those concerned with implicit configural learning (Geyer, Baumgartner, Müller, & Pollmann, 2012; Preston & Gabrieli, 2008; Westerberg, Miller, Reber, Cohen, & Paller, 2011). Along these lines, we suggest that recognition and visual search are supported by different memory processes. Further, the dissociation between the transfer of contextual cueing (Experiments 3, 4) and explicit recognition (Experiments 1–4) suggests that the memory underlying explicit learning is more flexible than that underlying implicit configural learning.

Conclusion

In sum, our experiments suggest that when display orientation switches and icons are rearranged, the traditional position-order remapping method used in current mobile touch devices is suboptimal in aiding search performance. Comparing and contrasting three alternative methods of icons remapping, we found that when using local-invariant or central-invariant remapping, contextual cueing continues to enhance (target) icon localization performance. Although the global-rotation method may be intuitive for users, it might introduce additional mental-rotation processes that are detrimental to localization performance. Our findings thus have implications for alternative interface design guidelines for icon rearrangement in mobile devices. Open questions awaiting further research concern how to optimize local invariance regions and what the effects of active manual rotation are.

Keywords: contextual cueing, visual search, mobile interface, icon remapping

Acknowledgments

This work is supported in part by German Research council (DFG) project grant SH166 to Z. S. and by Chinese Scholarship Council (CSC) to X. Z. and L. J.

Commercial relationships: none.

Corresponding author: Zhuanghua Shi.
Email: shi@psy.lmu.de.
Address: Experimental Psychology, Department Psychology, Ludwig-Maximilians-Universität München, 80802, Munich, Germany.

Footnote

[1] All icons were selected from www.softicons.com, available under a Creative Commons Attribution license.

References

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Giving a helping hand: Effects of joint attention on mental rotation of body parts. *Experimental Brain Research, 211,* 531–545. doi:10.1007/s00221-011-2625-z.

Borst, G., Kievit, R. A., Thompson, W. L., & Kosslyn, S. M. (2011). Mental rotation is not easily cognitively penetrable. *Journal of Cognitive Psychology, 23,* 60–75.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 798–815. doi:10.1037/0096-1523.33.4.798.

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

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,* 699–706. doi:10.1037/0278-7393.32.4.699.

Chua, K. P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics, 65,* 72–80.

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

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. doi:10.1006/cogp.1998.0681.

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

Chun, M. M., & Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition, 7,* 65–81. doi:10.1080/135062800394685.

Conci, M., & Müller, H. J. (2012). Contextual learning of multiple target locations in visual search. *Visual Cognition, 20*, 746–770. doi:10.1080/13506285.2012.694376.

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

Conci, M., & von Mühlenen, A. (2009). Region segmentation and contextual cuing in visual search. [Research Support, Non-U.S. Gov't]. *Attention, Perception & Psychophysics, 71*, 1514–1524. doi:10.3758/APP.71.7.1514.

Conci, M., & von Mühlenen, A. (2011). Limitations of perceptual segmentation on contextual cueing in visual search. *Visual Cognition, 19*, 203–233.

Geyer, T., Baumgartner, F. Müller, H. J., & Pollmann, S. (2012). Medial temporal lobe-dependent repetition suppression and enhancement due to implicit versus explicit processing of individual repeated search displays. *Frontiers in Human Neuroscience, 6*(272), doi:10.3389/fnhum.2012.00272.

Geyer, T., Shi, Z., & Müller, H. J. (2010). Contextual cueing in multiconjunction visual search is dependent on color- and configuration-based intertrial contingencies. [Research Support, Non-U.S. Gov't]. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 515–532. doi:10.1037/a0017448.

Ionta, S., & Blanke, O. (2009). Differential influence of hands posture on mental rotation of hands and feet in left and right handers. *Experimental Brain Research, 195*, 207–217. doi:10.1007/s00221-009-1770-0.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review, 12,* 100–106.

Jiang, Y., Song, J. H., & Rigas, A. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review, 12,* 524–529.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—Configuration or individual locations? *Perception and Psychophysics, 66,* 454–463.

Kunar, M. A., Flusberg, S., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cueing guide the deployment of attention. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 816–828.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cueing by global features. *Perception & Psychophysics, 68,* 1204–1216.

Makovski, T., & Jiang, Y. (2010). Contextual cost: When a visual-search target is not where it should be. [Research Support, N.I.H., Extramural]. *Quarterly Journal of Experimental Psychology, 63*, 216–225. doi:10.1080/17470210903281590.

Manginelli, A. A., & Pollmann, S. (2009). Misleading contextual cues: How do they affect visual search? *Psychological Research, 73*, 212–221. doi:10.1007/s00426-008-0211-1.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9,* 273–302.

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

Preston, A. R., & Gabrieli, J. D. E. Dissociation between explicit memory and configural memory in the human medial temporal lobe. *Cerebral Cortex, 18*, 2192–2207.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, New Series, 171,* 701–703.

Shomstein, S., & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. *Psychonomic Bulletin & Review, 11,* 247–253.

Song, J. H., & Jiang, Y. (2005). Connecting the past with the present: How do humans match an incoming visual display with visual memory? *Journal of Vision, 5*(4):4, 322–330, http://www.journalofvision.org/content/5/4/4, doi:10.1167/5.4.4. [PubMed] [Article]

Westerberg, C. E., Miller, B. B., Reber, P. J., Cohen, N. J., & Paller, K. A. (2011). Neural correlates of contextual cueing are modulated by explicit learning. *Neuropsychologia, 49*, 3439–3447.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 693–716.

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2011). Two (or three) is one too many: Testing the flexibility of contextual cueing with multiple target locations. [Research Support, Non-U.S. Gov't]. *Attention, Perception & Psychophysics, 73*, 2065–2076. doi:10.3758/s13414-011-0175-x.