Role of foveal and peripheral vision in contextual cueing and its retrieval in real-world scenes

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

Dr. rer. nat.

genehmigt durch die Fakultät für Naturwissenschaften der Otto-von-Guericke- Universität Magdeburg

von M.Sc., Eleonora Porracin

geb. am 19. September 1988 in Pordenone (Italien)

Gutacher: Prof. Dr. Stefan Pollmann

Prof. Dr. George Mather

Eingereicht am 25. April 2018

Verteidigt am 28. November 2018

To my parents

SUMMARY

Efficient visual search requires the selection of goal-relevant information in the environment. This mechanism is subject to visual cues that attract the observer's attention. In complex real-world environments, contextual information is a powerful tool for guiding visual attention. This idea was first tested in a paradigm called *contextual cueing*. Contextual cueing is a form of learning in which the repeated exposure to a specific arrangement of target and distractors leads to a progressively more efficient visual search (Chun & Jiang, 1998). The memory for contextual information guides visual attention towards task-relevant parts of a scene, enhancing the efficiency in visual search tasks. Therefore, search time in repeated scenes decreases across repetitions, while the search time for new scenes remains constant throughout the experiment. The benefits in the repeated condition arise from the memory for scene-target covariation, explicitly encoded in visual long-term memory. Participants showed to retrieve the exact position of the target in repeated configurations. The contextual cueing effect is particularly strong in naturalistic environments because a scene is a semantically coherent view of a real-world environment with a background and multiple elements arranged in a consistent manner. The scene itself intrinsically suggests the range of plausible objects that can occur in a particular situation and the position the objects occupy in relation to each other. In real-world scenes, the context is highly predictable and the visual system takes advantage of this predictability. Contextual cuing is considered a form of spatial statistical learning. For the

learning of the contingencies to occur, the visual system matches the incoming display with previous acquired memories: the easier the match, the faster the search for relevant targets. In the past decade, two contradictory theories have been debated regarding which part of the display is the most crucial, in order to detect contextual regularities.

In the current work, a gaze-contingent technique was applied to fully dissociate the central/local context from the peripheral/global context and thereby solve the question of global versus local dominance. If contextual cueing in real-world scenes is based on local context, foveal vision should be sufficient for the effect to developed with a gaze-contingent window. At the opposite, if global context is necessary to guide visual search, exploring the scene with a simulated or natural scotoma should not impair contextual cueing effects. In order to investigate contextual cueing in real-world scenes, a set of twelve different 3D-rendered images of indoor environments were created. Participants were tested in a visual search task under different viewing conditions and subsequently, they performed a recognition task, where they had to retrieve the position of the target in the repeated displays, in order to establish the implicit or explicit nature of the memory in contextual cueing. In Experiment 1 contextual effects in real-world scenes were examined under tunnel vision. That is, participants could rely only on central vision while performing the search. In Experiment 2, the reverse condition was investigated, to determine the contribution of peripheral vision. Participants searched for the target in real-world scenes with a gaze-contingent scotoma. Finally, in Experiment 3 contextual cueing was inspected in patients suffering from age-related macular degeneration (AMD), in which the central part of the visual field is progressively lost, to determine whether realistic scenes would help patients with AMD to efficiently learn spatial regularities and facilitate search for a target. The results showed that search was faster in the repeated configurations compared to the configurations in which the target was in a new position at each presentation. Thus, contextual cueing developed over the course of the experiment when the search was performed with impaired viewing conditions. The explicit memory representation of the scenes interacted with visual attention to enable efficient search in repeated configurations.

Although the visual system adapted to be sensitive to repeated spatial locations, contextual learning is encoded with respect to a hierarchical representation of the scene, that is, memory for the target location is learnt relative to optimal sub-regions of the scene. The present results suggest, that visual search is optimally driven by the local context around the target. However, when the local information is not available, learning of the repeated target position can also take place in periphery. Furthermore, patients with AMD develop an eccentric extrafoveal retinal spot outside the scotoma, called preferred retinal locus (PRL), that acts like a pseudo-fovea. The PRL allows patients to relocate visual attention to the peripheral part of the visual field for encoding the objects and their location and storing target position in visual long-term memory.

DEUTSCHE ZUSAMMENFASSUNG

Effiziente visuelle Suche erfordert die Auswahl von zielrelevanten Informationen in der Umgebung. Dieser Mechanismus unterliegt visuellen Hinweisen, die die Aufmerksamkeit des Betrachters auf sich ziehen. In komplexen realen Umgebungen sind kontextbezogene Informationen ein leistungsfähiges Werkzeug, um visuelle Aufmerksamkeit zu lenken. Diese Idee wurde zuerst in einem Paradigma namens Contextual Cueing getestet. Kontextuelles Cueing ist eine Form des Lernens, bei der die wiederholte Präsentation einer spezifischen Anordnung von Ziel- und Distraktoren zu einer zunehmend effizienteren visuellen Suche führt (Chun & Jiang, 1998). Der Speicher für kontextabhängige Informationen lenkt die visuelle Aufmerksamkeit auf aufgabenrelevante Teile einer Szene und erhöht die Effizienz visueller Suchaufgaben. Daher nimmt die Suchzeit in wiederholten Szenen über Wiederholungen hinweg ab, während die Suchzeit für neue Szenen während des gesamten Experiments konstant bleibt. Der kontextabhängige Cueing-Effekt ist in naturalistischen Umgebungen besonders stark, da eine Szene eine semantisch kohärentes Abbild einer realen Umgebung, mit auf konsistente Weise angeordnetem Hintergrund und Elementen, ist. Die Szene selbst suggeriert intrinsisch, wo verschiedene Elemente plausibel auftreten können und die Position, die die Objekte in Bezug zueinander einnehmen. In realen Szenen ist der Kontext daher einfacher vorhersehbar und das visuelle System nutzt diese Vorhersagbarkeit aus.

Kontextuelle Suche wird als eine Form des räumlichen statistischen Lernens betrachtet. Für das Erlernen der Kontingenzen passt das visuelle System die derzeitige Präsentation an die zuvor erworbenen Erinnerungen an: Je einfacher die Übereinstimmung ist, desto schneller ist die Suche nach relevanten Zielen. Im letzten Jahrzehnt entstanden zwei widersprüchliche Theorien, welcher Teil des Displays am wichtigsten ist, um kontextuelle Regelmäßigkeiten zu erkennen.

In der vorliegenden Arbeit wurde eine Blick-kontingente Technik angewandt, um den zentralen / lokalen Kontext vom peripheren / globalen Kontext vollständig zu trennen und damit die Frage der globalen gegenüber der lokalen Dominanz zu beantworten. Wenn das kontextuelles Cueing in realen Szenen auf lokalem Kontext basiert, sollte das foveale Sehen ausreichend sein, damit sich der Effekt mit einem Blick-kontingenten Fenster entwickeln kann. Wenn im Gegensatz dazu der globale Kontext für die visuelle Suche notwendig ist, sollte die Erkundung der Szene mit einem simulierten oder natürlichen Skotom die kontextabhängigen Cueing-Effekte nicht beeinträchtigen. Um das kontextuelle Cueing in realen Szenen zu untersuchen, wurde ein Set von zwölf verschiedenen 3D-gerenderten Bildern von Innenumgebungen erstellt. Die Teilnehmenden wurden in einer visuellen Suchaufgabe unter verschiedenen Betrachtungsbedingungen getestet. Anschließend führten sie eine Erkennungsaufgabe durch, bei der sie die Position des Ziels in den wiederholten Anzeigen abrufen mussten, um die implizite oder explizite Natur des kontextabhängigen Speichers kontextabhängig zu ermitteln. In Experiment 1 wurden kontextuelle Effekte in realen Szenen unter Tunnelsicht untersucht. Das heißt, Teilnehmende konnten sich bei der Suche nur auf das zentrale Sehen verlassen. In Experiment 2 wurde der umgekehrte Zustand untersucht, um den Beitrag des peripheren Sehens zu bestimmen. Die Teilnehmende suchten in realen Szenen mit einem Blickkontingent-Skotom nach dem Ziel. Schließlich wurde in Experiment 3 kontextuelles Cueing bei Patientinnen und Patienten mit altersbedingter Makuladegeneration

(AMD) untersucht, um festzustellen, ob realistische Szenen diesen helfen würden, räumliche Gesetzmäßigkeiten effizient zu lernen und die Suche nach einem Ziel zu erleichtern. Die Ergebnisse zeigten, dass die Suche in den wiederholten Konfigurationen schneller war als in den Konfigurationen, in denen sich der Zielreiz bei jeder Präsentation in einer neuen Position befand. Das heißt, es entwickelte sich im Verlauf des Experiments kontextuelles Cueing, selbst wenn die Suche mit beeinträchtigten Betrachtungsbedingungen durchgeführt wurde. Die explizite Erinnerungspräsentation der Szenen interagierte mit visueller Aufmerksamkeit, um eine effiziente Suche in wiederholten Konfigurationen zu ermöglichen.

Obwohl das visuelle System angepasst ist, um effizient auf wiederholte räumliche Darbietung zu reagieren, wird räumliches kontextuelles Lernen in Bezug auf eine hierarchische Szenendarstellung codiert, das heißt, ein Speicher für den Zielort wird relativ zu optimalen Teilbereichen der Szene gelernt. Die vorliegenden Ergebnisse legen nahe, dass die visuelle Suche optimal durch den lokalen Kontext um das Ziel herum gesteuert wird. Wenn jedoch die lokale Information nicht verfügbar ist, kann das Lernen der wiederholten Zielposition auch in der Peripherie stattfinden. Darüber hinaus entwickeln Patientinnen und Patienten mit AMD einen exzentrischen extrafovealen Netzhautfleck außerhalb des Skotoms, den so genannten Preferred Retinal Locus (PRL), der sich wie eine Pseudofovea verhält. Der PRL ermöglicht es an AMD Leidenten, die visuelle Aufmerksamkeit auf den peripheren Teil des Gesichtsfeldes zu verlagern, um die Objekte und ihren Ort zu codieren und die Zielposition im visuellen Langzeitgedächtnis zu speichern.

CONTENTS

1

INTRODUCTION

I. WHERE DO WE LOOK?

Vision is remarkable. Starting from the light that bounces into the eyes from the objects around us, we can gain an impression of the complex world that surrounds us. The human visual system starts with the eye, where the visual input is analyzed thanks to the interaction of different and yet complementary regions in the retina, the layer of tissue at the back of the eye. In the retina there are two different types of receptors for vision, called rods and cones. The visual receptors contain light-sensitive pigments that, reacting to the light, trigger an electrical signal that is transmitted to the brain. In the different regions of the retina the distribution of the receptors is distinctly non-uniform. In 1935, Osterberg was the first to measure the photoreceptor density and their retinal location. He quantified the highest density of cones in the fovea (roughly 6 million), with a rapid decrease in the cone density within several degrees from the foveal center. For rods (roughly 120 million), he observed a rod-free zone in the fovea and a rapid increase in density at approximately 20° eccentricity and an asymptote into the far periphery (Figure 1).

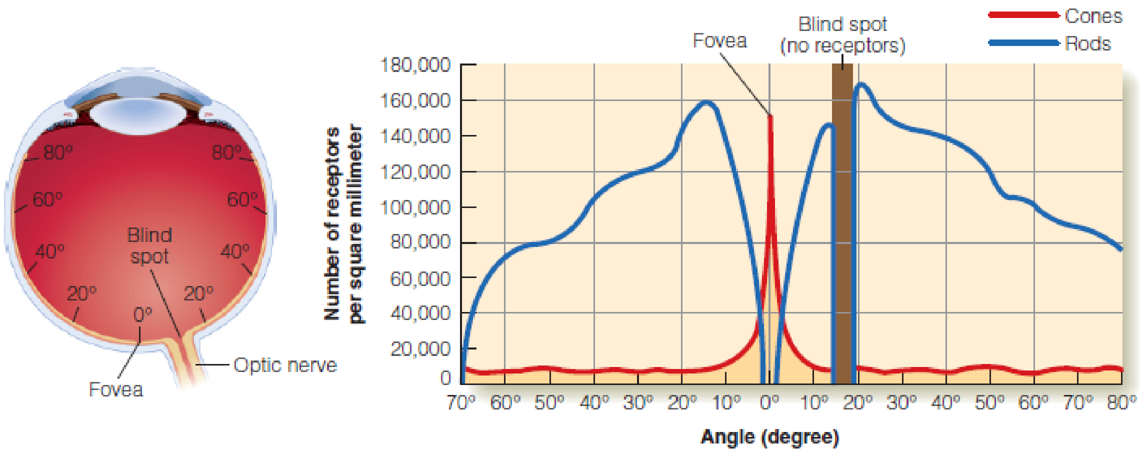


Figure 1. Distribution of rods and cones in the human retina. The graph illustrates that cones are present at a low density in the periphery of the retina, with a sharp peak in the center of the fovea. Conversely, rods are present at high density mostly at the periphery of the retina, with a rod-free zone in the fovea (*From Goldstein, E. B. (2010). Sensation and Perception 9th edition. Belmont (CA): Wadsworth, Cengage Learning, pp 28*).

The size of the visual field covers between 0° and approximately 70° eccentricity. Within this dimension it is possible to identify three major regions: fovea, parafovea and periphery (Polyak, 1941; Rodieck, 1998). The point with the highest visual acuity is the fovea (Figure 2), that, thanks to the densest concentration of cones, provides the best resolution power. The fovea is a tiny spatial region surrounding the center of the gaze (i.e. 0° eccentricity) and it extends for about 1° from the fixation point (Polyak, 1941; Yamada, 1969). The parafovea extends from the foveal rim for about 4° (Coletta & Williams, 1987; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). Vision scientists commonly refer to the part of the visual field beyond the parafoveal as periphery (Hollingworth, Schrock, & Henderson, 2001; Holmes, Cohen, Haith, & Morrison, 1977; Rayner et al., 1981; Shimozaki, Chen, Abbey, & Eckstein, 2007). The same classification is pursued in the present work.

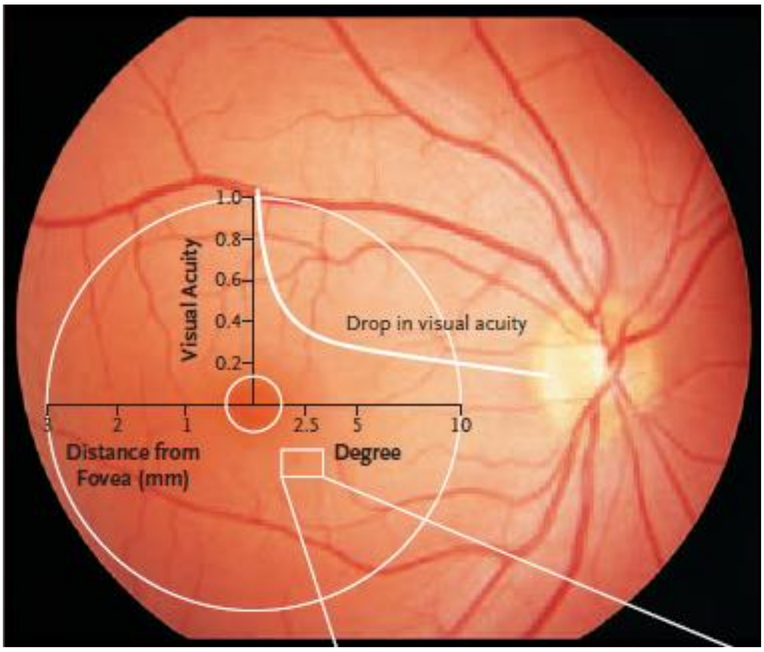


Figure 2. Fundus photograph of the retina of human eye. The area within the outer white circle (indicating a retinal diameter of about 6 mm) is the macula. The inner circle (diameter, 1.5 mm) borders the fovea, which is the central zone of the macula, where the preponderance of cones over rods is highest, therefore allowing the sharpest vision. The graph superimposed on this healthy retina shows the drop in attainable visual acuity in relation to the distance from the fovea (from De Jong, P. T. (2006). Age-related macular degeneration. *New England Journal of Medicine*, *355*(14), pp 1475).

The anatomically distinct regions of the retina are closely related to the way the two types of photoreceptors are distributed. Consequently, the structure of the receptors and their position on the retina has repercussions in the functional activity of the receptors, in particular for what concerns visual acuity, that is the ability to resolve small details (Campbell & Green, 1965; Green, 1970; Hirsch & Curcio, 1989; Hirsch & Miller, 1987; Miller, 1979; Snyder & Miller, 1977; Williams, 1985; 1986). There is a strong inverse relationship between the retinal eccentricity of the visual region and the quality of the visual information it provides, more specifically the maximal visual acuity is obtained in the fovea with only cones and it gradually diminishes towards the periphery (Yarbus, 1967; Figure 3).

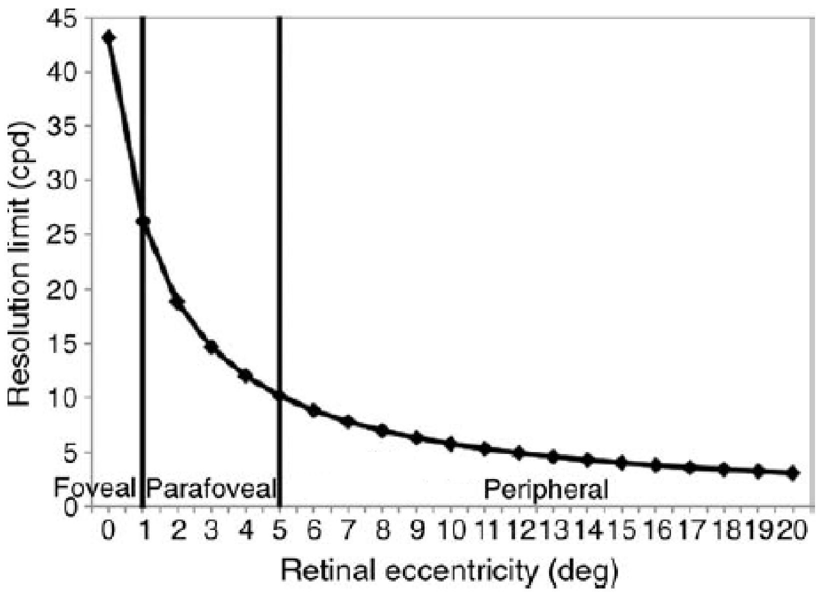


Figure 3. Limits of the visual resolution as a function of retinal eccentricity, with the visual field divided in three regions: fovea, parafovea and periphery (Adapted from Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), pp. 2).

Anatomically, the cones are found primarily in the central retina, within 10° of the fovea, their structure is smaller and they are tightly packed. Instead, the rods are bigger, widely spaced and spread mostly in the peripheral retina. More importantly, the cone's greater acuity is based on the convergence (i.e. the way the receptors send the signal to the ganglion cells). The difference between cones and rods is that a single cone sends the signal to a single ganglion cell directly, and therefore provides more detailed vision. Conversely, several rods converge to a single ganglion cell, this greater convergence translates into a greater sensitivity to dim light but ultimately to a poor acuity (Sterling, Freed, & Smith, 1988). To summarize, low convergence results in the high acuity we can find in cones (foveal vision), whereas, high convergence results in the poor acuity we can find in rods (peripheral vision).

Therefore, the way the cones and rods are organized in the retina influences what and how we perceive. High visual acuity is usually appreciated when we need to search for an object hidden among many others. In this case we tend to move the eyes for the purpose of scanning every single item. The reason why we adopt this behaviour is that we bring the object we want to inspect in the cone-rich part of the retina, the fovea, while the objects in the periphery are seen less clearly and blurred.

In order to process information from the visual field, we selectively explore the environment by moving the eyes, alternating periods of fixation to rapid changes of eye position, the saccades. Eye movements are the mechanism by which the gaze is brought to the content that needs to be inspected by the high-resolution foveal vision. The choice of where to direct the eyes is not random, but instead is influenced by properties of the scene as well as the goals of the observer (Yarbus, 1967). Thus, what we perceive and understand about the visual world is closely related to where our eyes are directed.

II. ATTENTION AND VISUAL SEARCH

The observer is actively involved in creating perception, because they direct their attention toward specific objects in the visual field and deliberately ignore others (Eriksen & Hoffman, 1973). The reason why we need to pay attention to some things and ignore others is that the perceptual system has limited capacity for processing information. Thus, the visual system needs to prioritize the processing of the information that in a given moment is more relevant (Pashler, Johnston & Ruthruff, 2001). This mechanism is called spatial visual attention. Allocating spatial visual attention to a specific part of the visual field allows the observer to more quickly and effectively process the attended element and its surroundings (Hoffman & Nelson, 1981). Visual spatial attention is often compared to a “spotlight” (Cave & Bichot, 1999) to emphasize the idea of a limited locus that moves across different locations. Although, spatial attention and eye movements are tightly linked (Deubel & Schneider, 1996), spatial attention can also be dissociated from eye movements. That is, it can be directed to the peripheral portion of the visual field by actively moving the eyes, so that the stimulus is brought into the fovea (overt attention) or the stimulus can be attended without being foveated (covert attention; Posner, 1980).

The visual environment is overcrowded with a large amount of stimuli competing for visual awareness. To deal with this overwhelming situation, visual attention must select and prioritize the processing of relevant information at any given time, ignoring the irrelevant stimuli (Chun & Wolfe, 2001). This mechanism is subject to visual cues that attract the observer's attention. Findings from visual search tasks suggest that visual attention can be captured by stimulus salience, namely the properties of the stimulus, such as luminance, movement and color (bottom-up factors). Or, visual attention can be strongly directed toward a particular object by cognitive factors, such as memory of previous experiences, goals, knowledge and task-demands (top-down factors. Bacon & Egeth, 1994; Henderson, Weeks, & Hollingworth, 1999; Theeuwes, 1992; Treisman, & Gelade, 1980; Wang, Cavanagh, & Green, 1994; Wolfe, 1994;1998; Yantis, 1998; Yantis & Hillstrom, 1994;). When in complex visual environments bottom-up factors are not sufficient (e.g. the middle of Piccadilly Circus in the rush-hour, when a red flickering light is not enough to capture the attention), a strong cue for visual attention is contextual information, that informs what object should appear in a certain context and where. Thus, contextual information, acquired with experience, can affect visual processing in a knowledge-based manner (top-down factors; Chun, 2000).

III. THE CONTEXTUAL CUEING PARADIGM: LEARNING SPATIAL CONTRAINTS IN VISUAL SEARCH

A scene is a semantically coherent human-scaled view of a real-world environment with a background and multiple elements arranged in a spatially licensed manner (Henderson &

Hollingworth, 1999). Real-world scenes differ from the stimuli commonly used in lab experiments (i.e. arrays of letters) because of their intrinsic semantic and syntactic constraints (Biederman, Mezzanotte, & Rabinowitz, 1982).

In scenes, context guides eye movements so that the observer processes the most relevant items in the visual field with foveal vision. This is demonstrated, for example, by the finding that people tend to fixate on the parts of the scene that are judged more informative (Yarbus, 1967). The reason why observers are inclined to focus their attention on certain spots, ignoring others, is that objects in a scene almost always occur in a rich visual context and this context is highly predictable (Biederman, 1972). Indeed, the scene itself intrinsically suggests the range of plausible objects that can be expected to occur in a particular situation (Palmer, 1975). Moreover, context suggests the position the objects occupy in a scene and the relation they have with respect to each other (Biederman et al., 1982; Hollingworth, 2006). For example, in entering a bedroom a person will probably expect to find a bed, on top of the bed a pillow and a blanket, the bedside table next to the bed, a closet, possibly a mirror. All the constraints the context provides are particularly helpful because they aid the observer's visual search and object recognition (Biederman, 1972; Palmer, 1975).

As the visual environment is stable in its structure, the visual system takes advantage of this predictability. In most scenes, the items covary with each other, so that finding an element predicts the presence of the other one. This repetitiveness, established over many times of encountering two elements in association, enables a schema. The schema is the memory representation of a scene type and contains information about the items and the spatial relations between the items that form that specific type (Henderson & Hollingworth, 1999). The visual system is sensitive to the regularities in the context and, hence, tend to remember them, in order to reduce the complexity of the visual environment by increasing its predictability (Gibson, 1963). Thereby, observers learn the regions of interest and the regularities in the visual input over time, based on the scenes' statistics (i.e. co-occurrence frequencies of elements in the visual world; Baker, Olson, & Behrmann, 2004; Fiser & Aslin, 2005). By encoding and retaining these regularities, observers can optimise their behaviour and increase their efficiency in visual processing (Chun & Turk-Browne, 2008).

The idea that the context provides a cue for the purpose of optimizing visual behaviour was first tested by Chun and Jiang (1998) in a paradigm called *spatial contextual cueing*. Contextual cueing is a form of learning where the repeated exposure to a specific arrangement of target and distractors leads to a progressively more efficient visual search (Chun & Jiang, 1998). With the contextual cueing paradigm, Chun and Jiang showed that implicit memory of contextual information guides visual attention towards task-relevant parts of a scene. In their experiment, participants search for a T (target) hidden among Ls (distractors). Critically, over the course of the experiment, half of the search displays are consistently repeated, so that the arrangement of target and distractors is kept constant. Whereas, in the other half, the target and the distractors change position randomly. The reaction times (RTs) during the search task for repeated displays are faster compared to those for new displays (Figure 4). Thus, the results prove that participants developed a sensitivity to the context, so that memory of the repeated configurations guides visual attention. Interestingly, several studies suggested that the knowledge underlying the contextual cueing effect is inaccessible to awareness, that is, implicit (Chun & Jiang, 1998; Chun & Jiang, 2003). In other words, the participants are not able to explicitly distinguish repeated configurations from the new ones and they cannot identify the target position in the repeated displays. Moreover, the effect proves to be long-lasting, since memory representation of spatial context persisted for at least one week, producing a significant effect even when participants were tested a second time after a week from the initial learning (Chun & Jiang, 2003; Jiang, Song, & Riga, 2005).

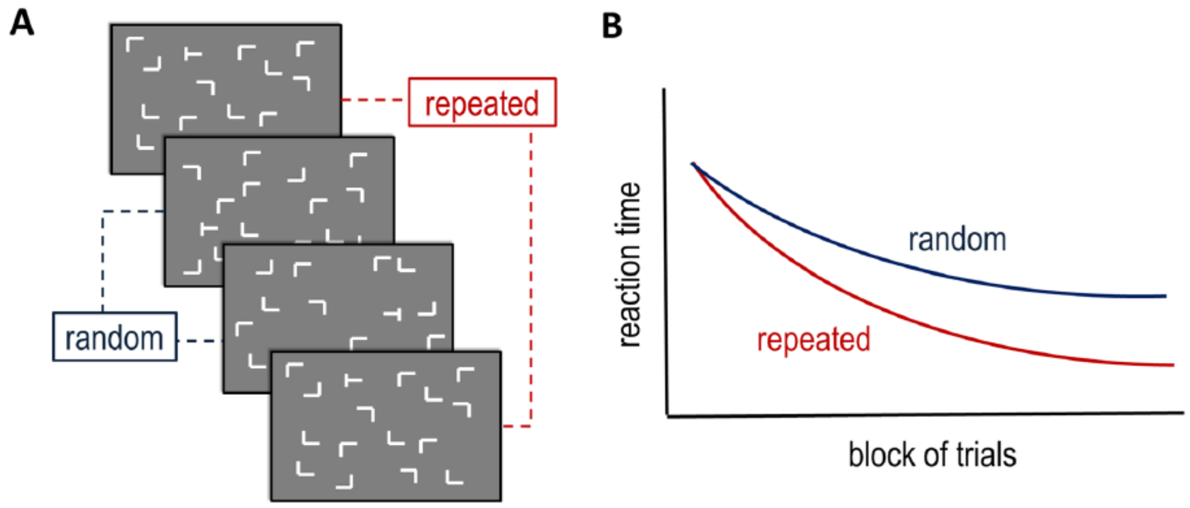


Figure 4. A) Example of a search display in the classic contextual cueing experiment. Participants are asked to search for the T among the Ls. Half of the search displays are repeated, whereas the other half are new. B) The typical pattern of results shows that participants become faster in finding the target in the repeated displays. (From Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2015). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review, 23(1)*, pp. 89).

In recent years the role of consciousness in contextual cueing has been debated. Although, learning of repeated configuration has been repeatedly demonstrated to be implicit in letter arrays (i.e. T and Ls; e.g. Chun & Jiang, 1998; 2003; Olson & Chun, 2002; Song & Jiang, 2005), it is now well established that cueing effects in real-world scenes are based on explicit memory (e.g. Brockmole, Castelhano, & Henderson, 2006; Brockmole & Henderson, 2006a; 2006b; Torralba, Oliva, Castelhano, & Henderson, 2006).

Contextual cueing in real-world scenes was studied at first in 2006 by Brockmole and Henderson. The idea raised from the assumption that although the stimuli used in the classic contextual cueing experiment follow the kind of structure available in real-world environments, an array of letters can lack in realism (Chun, 2003). The aim of the study was to investigate whether the observers are as sensitive to the repeated association scene-target position when viewing real-world environments. Participants were asked to search for a letter incorporated in full-colour photographs of real-world scenes (i.e. interior rooms). Like in the classic paradigm, two conditions were tested: repeated vs new. The repeated condition contained consistent association between the target's location and the scene, whereas in the new condition the scenes were presented only once during the experiment. Consistent with the contextual cueing effect, the search time in the new condition remained constant throughout the experiment. In contrast, search time in repeated scenes decreased across repetitions. Thus, participants proved to be sensitive to the covariation between target location and real-world scene. Moreover, the effect was achieved five times faster and led to a benefit twenty times greater in comparison with letter array displays (Brockmole & Henderson, 2006). These benefits can be ascribed to the fact that the memory for scene-target covariation was explicitly encoded. When asked to recognize the repeated displays, the participants performed above chance (97% hit). Additionally, the participants showed a remarkable accuracy in indicating the exact position of the target (with an error of 1.7 cm from the actual location). Therefore, this study proved that repeated exposure to the stimuli led to a more efficient visual search and allowed the observers to learn the location of the target in the associated scene, by creating an explicit memory, thanks to the semantic contribution of the scene content. The results were consistent with the idea that scenes can be quickly identified and categorized, enabling expectations about the components in the scene (Friedman, 1979; Potter, 1976), which consequently guide visual attention to informative parts in the scene.

IV. GLOBAL AND LOCAL CONTEXT

Contextual cuing is considered a form of spatial statistical learning (Chun & Turk-Browne, 2008; Goujon, Didierjean, & Thorpe, 2015; Zellin, von Mühlenen, Müller, & Conci, 2013). By encoding and maintaining in memory the regularities they repeatedly encounter in the environment, observers can reduce the cognitive complexity of the visual world, and consequently enhance visual search. In order to determine these statistical regularities, the observers identify the probabilistic associations between the configurations of objects (Baker et al., 2004; Fiser & Aslin, 2005). A crucial issue that has been debated in the past decade is which part of the display is the most valuable, in order to detect contextual regularities. In other words, does contextual cueing rely on the learning of the global configuration of the display, or the opposite, on the local context neighbouring the target?

Several studies showed that observers are not sensitive to entire display in the same way. Jiang and Wagner (2004) found evidence that the target location was learnt relative to the individual location of the objects, rather than the global configuration of the display, suggesting that learning could transfer to a new context as long as the some local aspects were preserved. Along the same line, Olson & Chun (2002) conducted an experiment aimed to investigate whether the global configuration is useful for locating the target or alternatively the spatial proximity modulates the contextual learning the most. In their paradigm the display was divided into two regions, the repeated context occupied one half of the screen, while the new context occupied the other half. The participants could not perceive the two parts of the display since no visual boundary was marked. The target appeared either within the repeated region (local or short-range condition) or in the new region (global or long-range condition). Contextual cueing was obtained only when the target was placed within the short-range condition, indicating that observers were more sensitive to the local area near the target. Therefore, spatial contextual cueing was encoded locally. The importance of local context was also confirmed by Brady and Chun (2007). In their experiment they further reduced the informative context to the only two (out of eleven) distractors contained in the same quadrant as the target, thus limiting the predictive information to 20% of the stimulus display. The results from both the model and the behavioural data, indicated that contextual cueing could be elicited from a remarkably minimal predictive information. Participants showed significant benefit for configurations in which only the target and the two closest distractors were repeated (with no significant difference than when the entire configuration was repeated), while no benefit was obtained from predictive information placed outside the target quadrant. Therefore, observers appeared to rely on local predictive information for the search. Based on the preferential encoding of the stimuli near the target, the authors stated that contextual cueing is based on the pairwise association between the target and the distractors immediately located around it. These findings are also consistent with the results of Song and Jiang (2005), in which successful retrieval of the learnt displays occurred when only 20-30% of the display matched the previous one. These pieces of evidence indicate that at the end of a successful search, the target is memorized together with its local context (Olson & Chun, 2002; Brady & Chun, 2007; Song & Jiang, 2005).

All the studies previously mentioned used letter arrays to prove that repeating only the local items near the target can be enough for developing contextual cueing. In contrast with those findings, Brockmole, Castelhano and Henderson (2006) demonstrated the importance of global context in visual search when contextual cueing is applied to real-world scenes. Their research aimed to test whether the target location is learnt given the local context surrounding the target or from the global context of the scene. Participants were presented with 3D-rendered illustrations of real-world environments (i.e. interior rooms) and asked to search for a target letter embedded in the scene. The scenes were manipulated in a way that in the learning phase the participants learnt the association between the scene and the target's position. In the second part of the experiment the local (i.e. the objects in close proximity to the target) and the global (i.e. the background scene) context were independently modified, so that half of the participants saw the modified local context in a familiar scene, while the other half saw the familiar local context in a different scene. Changes in the local information surrounding the target had no impact on contextual cueing, so that the search advantage in the repeated trials was preserved. However, changing the global scene completely eliminated the benefit of the contextual cueing in visual search. Furthermore, contextual cueing developed faster and yielded to greater effect when the global context was repeated compared to the local. Therefore, the results suggested that in realistic scenes observers associated the target with the global context. Similarly, Rosenbaum and Jiang (2013) found that predictive global context facilitated visual search. Participants were asked to search for a T among an array of Ls placed against a background scene of indoor or outdoor natural scenes. During the training phase the target location was predicted by both the background scene (i.e. the global context) and the array (i.e. the local context). Instead, in the testing phase the two types of clue were dissociated: participants were tested with displays that showed the same background context but a different array, or displays where the array was preserved but the background scene changed. The findings showed that the global context overshadowed the array context, thus significant contextual cueing was obtained when the scene in the background was the only predictive cue, but not when the array context was the only cue. The data suggested that scene and array-based cues did not simultaneously influence contextual cueing, but instead the stronger cue dominated the contextual learning. Specifically, the background global context drove spatial attention more efficiently to the target and therefore assumed an important role in contextual learning and its retrieval (Brockmole et al., 2006; Kunar, Flusberg, & Wolfe, 2006; Rosenbaum & Jiang, 2013).

V. CONTEXTUAL CUEING UNDER TUNNEL AND SCOTOMA VISION

The above presented studies bring a valuable contribution to determine the role of local and global context in contextual cueing and help to understand what guides visual search in familiar environments. However, global and local context might not be clearly dissociated in the paradigms used so far.

As previously reported, during visual search the eyes move in order to bring the regions of interest in the display into the fovea, the part of the retina with higher visual acuity and therefore the better resolution. Our visual system is divided into central and peripheral vision. Central vision concerns the information processed by the fovea and parafovea, peripheral vision covers the rest of the visual field. In everyday life, we examine the elements we directly look at in high resolution; in the meantime we still perceive information coming from the periphery, albeit at a lower resolution. For example, when driving a car we tend to keep our eyes on the street, looking directly at the car in front of us, nevertheless at the same time we perceive information coming from the side of the road, so that we can react in case, for example, a pedestrian suddenly crosses the street. In this respect, the local context can be operationalized as the information obtained from the central part of the visual field, while peripheral vision can be considered the global context. When locating a target in a scene, we tend to fixate the target and its surrounding with the central vision, while the rest of the scene falls into the peripheral portion of the visual field.

The relevance of peripheral information was established, among others, by van Asselen and Castelo-Branco (2009). In their experiment participants were tested in a classic contextual cueing task. Critically, the stimuli were placed around a fixation cross, that the participants were instructed to fixate. Without moving their eyes, they had to locate the target (the letter T among Ls) counting only on the information coming from peripheral vision. The results showed that when the configurations were repeated, search time was faster, indicating that contextual cueing took place. These findings are consistent with the idea that foveating the target is not necessary to guide attention to relevant targets during visual search.

The concept of "local" and "global" can be difficult to define, since what is considered as local in one circumstance, can be seen as global within another (Brooks, Rasmussen, & Hollingworth, 2010). In order to examine whether pure central/local information in absence of any peripheral/global information (and vice versa) would be sufficient to develop contextual cueing, a gaze-contingent technique may be applied to the classic contextual cueing paradigm. Based on real-time tracking of the eye's position, the gaze-contingent is continuously updated as a function of the participant's current gaze location (i.e. x and y coordinates recorded by the eye-tracking device). The gaze-contingent technique allows for the investigation of specific cognitive processes, controlling which information is fed to the visual system at any given time. The role of central versus peripheral vision in scene perception has therefore been studied using paradigms known as "window" and "scotoma". The window paradigm (McConkie & Rayner, 1975; Reder, 1973) gives the feeling of viewing the scene through a tunnel, so that only the part of the visual field around the gaze location (foveal information) is unaltered, while outside the window the image is absent or altered. The scotoma paradigm (Rayner & Bertera, 1979) is the reverse technique, the central information is blocked from the view, permitting only the peripheral information to be processed (Figure 5).

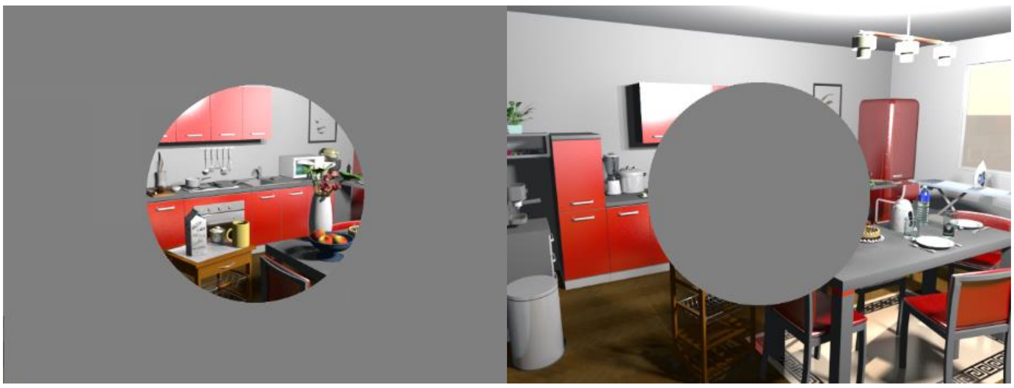


Figure 5. Examples of "window" and "scotoma".

If contextual cueing is based on foveal/local information, as suggested by Olson and Chun (2002) and by Brady and Chun (2007), the availability of local context should be sufficient to develop contextual cueing using a gaze-contingent window, in which only the items near the target are visible. On the contrary, if the contribution of the global background information is necessary (Brockmole et al., 2006; Rosenbaum & Jiang, 2013) for guiding visual search, the loss of foveal vision because of a simulated scotoma may not interfere with the learning of contextual regularities in the scene.

Recently, Zang and colleagues (Zang, Jia, Müller, & Shi, 2015) examined whether central vision, in absence of peripheral context, was sufficient to generate contextual cueing. In order to test the role of local context, a gaze-contingent window was simulated during a classic contextual cueing task (i.e. searching for the target T among distractors L). The experiment consisted of three sessions: the training session, where the search display was visible only through the gaze-contingent window; a transfer session, where the whole display was fully visible, in order to check if the context was learnt despite the tunnel vision; and at the end, a recognition session to test whether the subjects encoded the displays implicitly or explicitly. The size of the window was manipulated. In Experiment 1 the window was 8° in diameter and allowed a fully visible para/foveal area, while the remainder display was darkened, providing no information about the layout of the search display. On average, 2 or 3 out of 12 items were visible through the window simulation. The results showed that no contextual cueing emerged in the training session. Surprisingly, however, contextual facilitation was evident when the display was made fully visible, suggesting that the context was learnt during the search, but it could not be efficiently used to speed up the search behaviour. Therefore, in Experiment 2 further peripheral information was added, increasing the size of the window to 12°. On average, 4 or 5 out of 12 items were visible through the window. In this case, the results indicated a strong contextual cueing effect in both training and transfer sessions. Taken together, the pattern of findings suggested that although spatial properties could be acquired with a narrow gaze-contingent window, peripheral/global information was necessary for the cueing effect to become manifest. Being able to see only a restricted area around the fixation point was not sufficient to efficiently guide visual search toward the target (Zang et al., 2015). A similar experiment was carried out by Geringswald and Pollmann (2015) with the purpose of investigating the impact of peripheral vision loss and central vision loss in learning of spatial regularities during a classical contextual cueing task. Participants performed a visual search task in which they were asked to find the target T among Ls while viewing the stimulus through a gaze-contingent display, either a window or a scotoma. The simulation was 9° in diameter, therefore involving on average 3.36 search items. The experiment was divided into three phases: the learning phase, where participants completed the search with the gaze-contingent display; the test phase, where the viewing restriction was removed and a final recognition test. When the visual search task was performed under tunnel vision, that is, with a gaze-contingent window that prevented the processing of peripheral information, contextual cueing failed to occur, impeding the learning of the repeated configurations. Moreover, no facilitation was observed in the test phase, when the window was removed. Tunnel vision prevented the learning itself of the repeated configurations. Thus, the findings further emphasize the important role that global context plays in contextual cueing. In the reverse situation, that is when the visual search was performed with the simulated scotoma, that prevented the foveation of the target and its surrounding, no contextual effect emerged during the learning phase. However, as opposed to the tunnel vision, a reliable benefit for the repeated configurations appeared when the scotoma was removed. The results suggested that the learning of spatial regularities occurred despite the scotoma, but the learnt configurations could not be used to efficiently guide visual search as long as the scotoma was present. In summary, these findings uncover the importance of peripheral information by the fact that learning of target-distractor association is not dependent on the foveation of the target and the items in its proximity, but on the contrary vanished without peripheral vision.

VI. NATURAL SCOTOMA VISION: AMD PATIENTS

The simulated scotoma realized with the gaze-contingent display resembles a clinical condition called age-related macular degeneration (AMD), in which patients progressively lose sight in the central part of the visual field.

AMD typically occurs in elderly people, in particular over 50 years of age, and represents the leading cause of blindness (Velez-Montoya, Oliver, Olson, Fine, Quiroz-Mercado, & Mandava, 2014). One of the first signs of AMD is the presence of yellow subretinal deposits, called drusen. The excess of these deposits progressively damage the retinal pigment epithelium of the macula. Affected areas lose visual function since the damage of the epithelium is associated with fallout of the photoreceptors. The early stage of the pathology is usually asymptomatic, so people still have normal vision, even though they might experience blurred vision, decreased contrast sensitivity and abnormal dark adaptation. Over time it is possible to detect an increase in the amount and the size of the drusen, spreading from the fovea to the periphery. At its late stage, AMD patients manifest severe visual loss, with central and pericentral scotomas (Figure 6). Late AMD is divided into two forms: “dry” and “wet”. Dry AMD (also called geographic atrophy) starts with a round hypopigmented spot at the center of the macula, vision loss is generally gradual. The symptoms are usually gaps in the image, distortion in the perception of straight lines, blurred vision. Wet AMD (choroidal neovascularization) is caused by hemorrhagic fluid leakage below the macula, vision loss occurs faster and suddenly (Figure 7; de Jong, 2006; Jager, Mieler, & Miller, 2008; Ambati & Fowler, 2012).

Figure 6. Simulation of age-related macular degeneration (AMD) vision in comparison to healthy vision (*From https://nei.nih.gov/health/examples*).

Figure 7. Fundus Photographs: A) The ocular fundus of a healthy eye, showing normal pigmentation and retinal blood vessels. (B) Late-stage AMD “dry” form, with regions of hypopigmentation in the macula, (C) Late-stage AMD “wet”, with leakage of blood below the macula (From Ambati, J., & Fowler, B. J. (2012). Mechanisms of age-related macular degeneration. *Neuron*, *75*(1), pp. 27).

AMD affects all functions of central vision: acuity, color vision, contrast sensitivity and high spatial resolution (Midena, Degli Angeli, Blarzino, Valenti, & Segato, 1997; Sjostrand & Friseu, 1977). High spatial frequencies are crucial in visual acuity, as they provide detailed information about contours and shapes. Because of the nature of the disorder, the performance of patients with AMD has been examined in several tasks that would normally require central vision. Investigating the repercussion of compromised foveal vision in patients can be more ecologic, than with a gaze-contingent simulated scotoma. Following this reasoning, Geringswald and colleagues (2013) investigated whether patients with AMD showed impaired contextual cueing, given their inability to perceive the foveal part of the visual field. The search behaviour of AMD patients, as well as that of age-matched healthy controls, was examined by using the classic contextual cueing paradigm (i.e. searching for a T among Ls). Contextual cueing was found to be preserved in healthy aging, despite slower responding time (Howard, Howard, Dennis, Yankovich, & Vaidya, 2004; Merrill, Conners, Roskos, Klinger, & Klinger, 2013), however, patients suffering from AMD showed a reduced search facilitation in repeated displays. In patients, the impairment was revealed in both the search time and the efficiency of the scan path. Thus, the loss of foveal vision impaired the efficiency in guiding the attention to the target in familiar displays (Geringswald, Herbik, Hoffmann, & Pollmann, 2013).

Previous research on AMD has focused on investigations with abstract stimuli, such as gratings, shapes and letters (Legge, Ross, Isenberg, & Lamay, 1992; Legge, Rubin, Pelli, & Schleske, 1985; Wang, Wilson, Locke, & Edwards, 2002). However, in everyday life we are in contact with complex environments and dynamic scenes, therefore understanding how central vision loss can be compensated in rich environments is crucial for the development of strategies that can efficiently benefit AMD patients. For instance, Tran and colleagues (2010) examined the contribution of peripheral vision in scene gist recognition. AMD patients were asked to categorize photographs of natural scenes divided into two main categories: nature/urban and indoor/outdoor. Participant were given a target (e.g. urban scenes) and asked to respond when they saw a photograph corresponding to the target. The results showed that AMD patients were able to perform the task with a high hit rate (around 80%). Scene recognition can therefore be accomplished with low-resolution peripheral vision, possibly because the global properties of the scene can compensate for the impaired ability of object identification (Tran, Rambaud, Despretz, & Boucart, 2010). Color perception is another important point of interest for research on central vision loss, since color vision is prerogative of the fovea, where the density of the cones is the highest. Surprisingly, patients with central vision loss exhibited greater benefit for colored objects in object recognition. When tested in a recognition task, AMD patients and a group of age-matched controls had to discriminate the presence of an animal or a face in a photograph. Face and animals could be either colored or gray. While the age-matched controls with intact vision were not affected by the manipulation, the AMD patients benefited significantly from the colored object for both faces and animals (Boucart, Despretz, Hladiuk, & Desmettre, 2008). Thus, color might facilitate the segmentation of the scene into its constituent parts, helping to define spatial contours (Fine, MacLeod, & Boynton, 2003). Low vision also affects sensitivity to high spatial frequencies, therefore facial expression recognition is a difficult task for patients with AMD. Nevertheless, Boucart and colleagues (2008) demonstrated that AMD patients failed in judging the strength of the emotion, but were able to rapidly categorized facial expressions, discriminating whether the expression was positive, negative or neutral. The authors argued that likely AMD patients based their decision on low spatial frequencies information, therefore taking into account the interaction between facial features, instead of the facial features per se (Boucart, Dinon, Despretz, Desmettre, Hladiuk, & Oliva, 2008).

Visual search is a common daily activity. Therefore, being able to remember regularities in the environment and, consequently, to guide attention to the salient parts of the visual field is important for adaptation to the environment and management of the resources. Given the efficiency AMD patients showed in using global regularities in complex scenes, it is still unsolved whether contextual information can efficiently guide visual search in real-world environments under foveal vision loss. Compared to artificial displays, searching in semantically rich and coherent scenes, where the objects are arranged in a meaningful way, might be easier, as patients can adopt exploration strategies they use in everyday life.

VII. OUTLINE OF THE EXPERIMENTS

The present dissertation aims to investigate the role of foveal and peripheral information in contextual cueing when viewing real-world scenes. Previous research has studied contextual effects in repeated displays, both in letter arrays and real-world scenes. In order to estimate the influence of foveal/local and global/peripheral information in visual search, the gaze-contingent technique was applied during the viewing of the displays. The displays consisted of naturalistic scenes. A set of twelve different 3D-rendered images of real-world environments were created (see Appendix) with the purpose of, on the one hand, enhancing the level of realism, on the other, controlling the manipulation of the objects in the scene.

In Experiment 1 (Chapter 2), I examined contextual effects in real-world scenes under tunnel vision. Specifically, whether having access only to the local information surrounding the target can speed up the visual search in repeated displays. Due to the gaze-contingent window, participants could rely only on central vision while performing the search. The results were compared with those obtained from a control group, which completed the task without any viewing restriction. Subsequently, participants from both groups were tested in a recognition task, where they had to recall the position of the target in the repeated displays. This was done to test whether the memory mechanism that drives contextual cueing was implicit or explicit.

In Experiment 2 (Chapter 3), the reverse condition was investigated. Participants searched for the target with a gaze-contingent scotoma. The goal was to test whether, in naturalistic scenes, peripheral vision can efficiently guide visual attention to previously learnt salient regions of the display, so that visual search would benefit from the repeated exposure. Additionally, it was explored whether the performance would be different than searching with full vision. Afterwards, the memory for the position of the target in repeated displays was analyzed, in order to determine whether contextual cueing rely on explicit or implicit memory of the repeated configurations.

In Experiment 3 (Chapter 4), contextual cueing was inspected in patients with AMD, that is, patients suffering from natural central vision loss. Using a natural condition in which foveal vision was blocked, was useful to clarify aspects of visual search that could have been jeopardized by the artificial scotoma. The purpose was, moreover, verifying whether using realistic scenes would help patients with AMD to efficiently learn spatial regularities in order to facilitate the search for a target. The same search task was proposed to an age-matched control group with intact vision. Furthermore, all participants completed a recognition task, in which they indicated the position the target occupied in the repeated scenes. The goal of the recognition test was to determine whether the participants became explicitly aware of the association scene-target position.

Finally, in the last section (Chapter 5) the results of the experiments are interpreted and discussed in relation to each other. The conclusions are discussed in a broader context, in relation to the current literature. Limitations are pointed out and possible applications are suggested.

2

CONTEXTUAL CUEING IN REAL-WORLD SCENES WITH SIMULATED WINDOW

The visual world is a complex environment, crowded with an overwhelming amount of objects and many of those might not be useful for a given task. Therefore, observers need to decide where to focus their visual awareness, in order to find what they are looking for and to ignore irrelevant stimuli (Chun & Wolfe, 2001). Visual search is the basic mechanism that allows us to select the most salient parts of the visual field and it is based on visual attention, which helps to identify and retain specific items and therefore prioritise the processing of relevant information over irrelevant (Eriksen & Hoffman, 1973).

An important cue for visual attention is contextual information. Contextual information informs the observers of which objects could appear in a scene and their position with respect to each other. Indeed, context tends to be highly predictable, because the real world is configured in a structured and stable manner (Henderson & Hollingworth, 1999). The visual system takes advantage of this predictability and creates associations between objects and the context they are usually found in (Chun, 2000).

Visual sensitivity to context has been demonstrated in a paradigm known as contextual cueing (Chun & Jiang, 1998;1999; Chun, 2000). Participants are asked to find a target (i.e. the letter T) among a set of distractors (i.e. the letter L). Critically, predictable configurations (i.e. where the arrangement of target and distractors is fixed) are repeated throughout the experiment, intermixed among new displays. The repeated arrangement of target and distractors cues the target position and leads to an improved search performance, as demonstrated by the faster reaction times in the repeated configurations compared to the new ones. This effect proves that contextual learning takes place across repetitions and facilitates visual search, by optimizing the localization of the relevant target. Interestingly, in symbolic configurations the learning is implicit, as observers are not aware of the repetitions and perform at chance level when asked to discriminate the repeated configurations from those they saw once (Chun & Jiang, 1998; 2003).

Further research showed that observers are not sensitive to the entire display in the same way. Olson and Chun (2002) demonstrated that attention is focused strongly on the elements near the target. Observers showed contextual cueing only in a short-range predictive condition (i.e. the local area neighboring the target). Analogously, Jiang and Wanger (2004) found that the target location was learnt relative to the individual location of the contextual item, rather than relative to the global configuration of the display, suggesting the prominent role of local context in contextual cueing. The results were confirmed by Brady and Chun (2007), both the computational model they developed and the behavioural data, suggested that observers preferentially encode the spatial configuration of elements localized near the target. Furthermore, contextual learning took place when only 20-30% of the display was repeated (Song & Jiang, 2005) and the distractors close to the target appeared to be more important than those far away (Brooks, Rasmussen, & Hollingworth, 2008). This evidence indicates that local context contributes most to contextual cueing (Olson & Chun, 2002; Brady & Chun, 2007; Brooks et al., 2008).

Opposite results were observed investigating contextual cueing under tunnel vision simulation (Geringswald & Pollmann, 2015), limiting the visible part of the display around fixation and excluding the global context. Visual search was severely impaired by gaze-contingent window, resulting in a lack of benefit for the repeated exposure to the old displays. During the visual search the repeated configurations of target and distractors did not provide any advantage in guiding efficiently visual search to the target. Moreover, contextual cueing was not observed even when the peripheral gaze-contingent was removed, indicating that the tunnel vision prevented the learning itself of the repeated configurations. Similar results were found with gaze-contingent window simulation that left a narrow visible area around the fixation (Zang, Jia, Müller, & Shi, 2015). No contextual cueing emerged in the visual search phase in which the visible window was 8° in diameter, yet the effect manifested afterwards, once the gaze-contingent window was removed and the display was made fully visible. However, when the tunnel area was extended to 12° in order to include additional global information, observers showed the learning effect proper of contextual cueing. Therefore, robust facilitation for the repeated displays versus the new ones occurred only when the visible area was wider, in order to include more additional, peripheral information. Taken together, the findings of this study suggested that although context could be acquired with a tunnel view of two or three items, the learnt context could not be used to effectively guide the search behaviour. The efficient learning of target-distractors association was obtained only when four of five elements were visible, providing a more exhaustive knowledge of the display. Indeed, the wider tunnel allowed the retrieval of the learning context, therefore a contextual cueing effect. Thus, these finding brought evidence that the global context played an important role in contextual learning. As previously reported (Chun & Jiang, 1998; Olson & Chun, 2002), the learning of the spatial constraints was implicit.

Outside the lab, objects are rarely perceived in an homogeneous array of elements, but they mostly occur in a rich and complex context (Biederman, 1972). Starting from this idea, Brockmole and Henderson (2006) investigated spatial context learning in real-world environments (i.e. indoor rooms). Real-world environments have, more strongly than letter arrays, a stable structure, with semantic constraints, where elements covary with each other in a predictable way. With the purpose of examining how regularities in real-world scenes guide visual attention, observers were asked to search for an arbitrary letter incorporated in photographs of indoor rooms. Like in the classic paradigm, half of the scenes were repeated, while the other half were new. Consistently with contextual cueing effect, observers showed to be sensitive to the relation between target position and scene, hence the search time required to find the target decreased across repetitions for the repeated scenes, although it remained constant in the novel scenes. This result proves that regularities in naturalistic scenes guide attention to the relevant position of the display, and therefore the search for the target becomes more efficient. Quite the contrary to previous results (Chun & Jiang, 1998; Olson & Chun, 2002), with real-world scenes, observers proved to remember explicitly the association target-scene, so that they recognized the repeated displays more often than those presented once and recalled accurately the correct target position.

Following up on these results, Brockmole, Castelhano and Henderson (2006) investigated whether contextual cueing in real-world scenes is associated with the local context neighbouring the target or, the opposite, with the global context of the scene. Observers were asked to search for an arbitrary letter within 3D rendered scenes of indoor rooms. Critically, in the second phase of the experiment the local (i.e. the group of objects in close proximity to the target) and the global context (i.e. the scene in background) were independently modified, in a way that half of the participants saw the altered local context while preserving the global context, whereas the other half saw the changed global context while maintaining the local context. After the observers learnt the location of the target in the scenes, changes in the local context had no consequences on contextual cueing, however, changes in the global context completely eliminated the benefit of the repeated target location. Moreover, the scene-target association was learnt faster and better when the global context was repeated across trials but not when only the local context was repeated. Similarly, Rosenbaum and Jiang (2013) found that predictive global context facilitated visual search. The background global context overshadowed the local context in efficiently driving spatial attention to the target, therefore proving to be the strongest cue in contextual learning and its retrieval. Thus, the combination of these findings suggests that in naturalistic scenes, observers associate the target position with the global context.

The purpose of the present study is to examine whether an efficient visual search in naturalistic scenes can be performed with a gaze-contingent tunnel window (Reder, 1973; McConkie & Rayner, 1975), that leaves visible only the foveal and parafoveal part of the visual field, therefore just the elements near the fixation. If contextual cueing in real-world scenes is based on local context, the foveal information should be sufficient for the effect to manifest with a simulated tunnel window. By contrast, if the global context is needed, the lack of peripheral information should prevent contextual cueing.

2.1. METHOD

2.1.1. Participants

Tunnel vision group. Twenty (8 males, 12 females; average age 23.6 , SD= 6.04) undergraduate students of the Otto-von-Guericke University of Magdeburg gave their consent to participate in the study.

Full vision group. Twenty (1 male, 19 females; average age 20.35, SD= 1.93) undergraduate students of the Otto-von-Guericke University of Magdeburg took part in the study as the control group.

All participants reported normal or corrected-to-normal visual acuity and were not informed about the purpose of the study. Written informed consent was provided before the beginning of the experiment, which was approved by the Ethics Committee of the University of Magdeburg. Participants could choose to be compensated with course credits or with a remuneration of 6 euros/hour.

2.1.2. Apparatus

The stimuli were presented and the responses recorded by using PsychoPy version 1.82 and a PC running Linux Debian. The monitor connected to the PC was a LCD BenQ XL2410T full color 24 inch HD, it was 521 mm (1920 pixels) wide and 293 mm (1080 pixels) high, with a refresh rate of 120 Hz. The stimuli were presented at a distance of 80 cm in a quiet and dimly lit room.

2.1.3. Stimuli

Stimuli consisted of 3D rendered illustrations of twelve indoor rooms. All the images were created with an open source interior design software (Sweet Home 3D, version 4.6, eTeks Paris, France. Available at http://www.sweethome3d.com. See Appendix) and represented a bathroom, a bedroom, a cinema room, a children's room, a garage, a recreation room, a kitchen, a library, a living room, a music room, an office and a study. The rooms were created balancing the colors of the environment and the amount of objects in the room, in order to avoid any pop-out effect. Each type of room was presented in a singular exemplar. The scenes were displayed at a resolution of 1200x900 pixels, subtending a visual angle of 22.83x17.12 degrees.

Each scene included a yellow cup, which constituted the target in the visual search task. The cup's position was equally allocated to one of the six equal-sized rectangular segments in which the scene was divided, three (left, middle, right) above respectively below the horizontal midline. The cup was always presented in a meaningful position (e.g. cup on a surface and not floating in the air) and clearly visible.

2.1.4. Design and procedure

The experiment was divided into two parts, separated by a short break.

Visual search. During the course of the first part of the experiment participants viewed 72 stimuli, divided into six blocks of twelve trials each. Each block contained six scenes in which the position of the target was repeated throughout the experiment, in the remaining six scenes the target was presented in a different segment at each block. The subjects were not instructed that in some scenes the target position was repeated during the experiment. Participants were asked to search for the yellow cup in each scene and to indicate the side of the handle by pressing the left or the right button of the mouse as quickly as possible. The side of the handle varied randomly, both in novel and repeated displays, so that no association of scene and response could be learnt across the repetitions. For each participant, the sequence of the stimuli was randomly assigned and the rooms that constituted the repeated displays were randomly drawn from the pool of all the rooms.

At the beginning of each trial, a fixation cross was presented at the centre of the screen for 1000 ms in black on a gray background. Participants were instructed to fixate on the cross. After the fixation cross, a blank gray screen was presented for 500 ms and after that, the search display was shown and remained visible until the response. Upon identifying the target, the subjects had to press the left button of the mouse when the cup had the handle on the left side or the right button of the mouse when the cup had the handle on the right side. After the response a feedback sound was provided (i.e. a high pitch tone for the correct response and a low pitch tone for the wrong response).

Participants in the experimental group (i.e. the tunnel group) viewed the scene through a gaze-contingent tunnel window (Reder, 1973; McConkie & Rayner, 1975). The tunnel window was 5° of visual angle in radius, covering the foveal and parafoveal region of the retina (Beard and Ahumada, 1999; Coletta and Williams, 1987; Larson and Loschky, 2009; Rayner et al. 1981). Throughout each trial, the position of the participant's left eye was recorded at the rate of 1000 Hz by an Eye-Link 1000 eye-tracker system (SR Research, Inc.), and according with the coordinates (x,y) provided by the eye-tracker device, the gaze-contingent window was drawn. When the gaze coordinates were unavailable due to eye blinks or signal losses, the screen was kept completely blank. Before each block a 9-point gaze calibration was performed, followed by a 9-point gaze validation at the beginning of each trial. The calibration and the validation procedure was repeated in case the error was bigger than 1° on average or bigger than 1.5° for the worse point. The tunnel simulation was not presented neither during the calibration nor the validation procedure (Figure 8).

Recognition task. In order to test whether the participants were able to explicitly indicate the position of the target in the repeated displays, they were examined in a recognition/memory task. Participants were not informed in advance about this second part of the experiment, in order to avoid any intentional memorization. After a short break, the subjects viewed the six real-world 3D rendered rooms that constituted the repeated condition in the main experiment. The scenes were rendered without the search target (i.e. the yellow cup) and participants were asked to move the mouse cursor to the location they remembered seeing the target. At the beginning of each trial, a fixation cross was presented at the centre of the gray background for 1000 ms and the scene was displayed until the response. No feedback was provided.

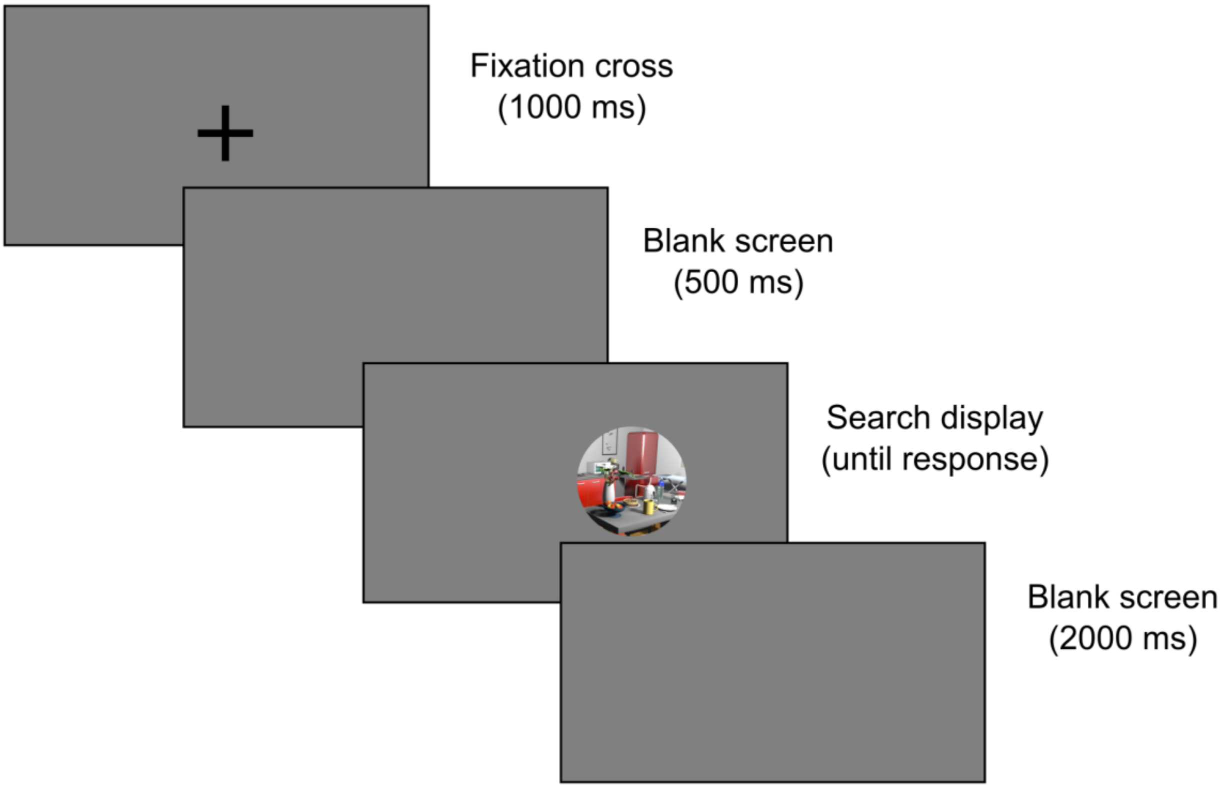


Figure 8. Example of a visual search trial with a gaze-contingent window.

2.2. RESULTS

2.2.1. Visual search.

Accuracy. In both groups the accuracy during the search task was very high. On average, in the experimental group the accuracy was 98.05%, whereas in the normal vision group it was 99.51%. In the tunnel vision group, one participant was excluded from further analysis due to the impossibility of tracking the gaze, therefore the gaze-contingent was not steadily drawn. This subject was replaced with a new one, who performed the exact same sequence. Only the correct responses were considered in the further analysis and outliers were excluded calculating the mean and two standard deviations as cut-off.

Response time. Only the correct responses were considered in the further analysis and the outliers were excluded when greater than two standard deviations from the mean computed on a subject-by-subject basis. For statistical analysis I performed a log transformation on the data and I ran a three-way mixed design ANOVA to calculate the contextual cueing effect, with Group (tunnel vision vs full vision) as between-subjects factor, Configuration (Repeated, Non-repeated) and Epoch (1-3) as within-subjects factors. With a total of six blocks, each epoch consisted of two blocks (Figure 9). A strong effect of group emerged ($F(1,38)=270.40$, $p< 0.001$), due to the longer search time required to complete the task with the gaze-contingent window. Response time, on average, was two times slower in the tunnel vision group (1966 ms) compared to the full vision group (1019 ms). Reaction times decreased over the course of the epochs, as reflected in the main effect of Epoch ($F(2,76)=27.96$, $p< 0.001$). A significant main effect of Configuration ($F(1,38)= 45.87$, $p< 0.001$) suggests that the search time was shorter in the repeated displays than in the non-repeated. Critically for the contextual cueing paradigm, the interaction of Epoch x Configuration was significant ($F(2,76)=3.68$, $p= 0.029$). The interaction demonstrates that the difference between repeated and non-repeated trials is larger at the end of the experiment, when compared to the beginning; search time decreased in the repeated condition from 2010 ms in the first epoch to 1069 ms in the last epoch, with a benefit of 320 ms, while the non-repeated condition was stable throughout the experiment. Furthermore, the two groups did not show a significant difference in the contextual cueing effect (Group x Epoch x Condition: $F(2,76)=0.19$, $p=0.82$).

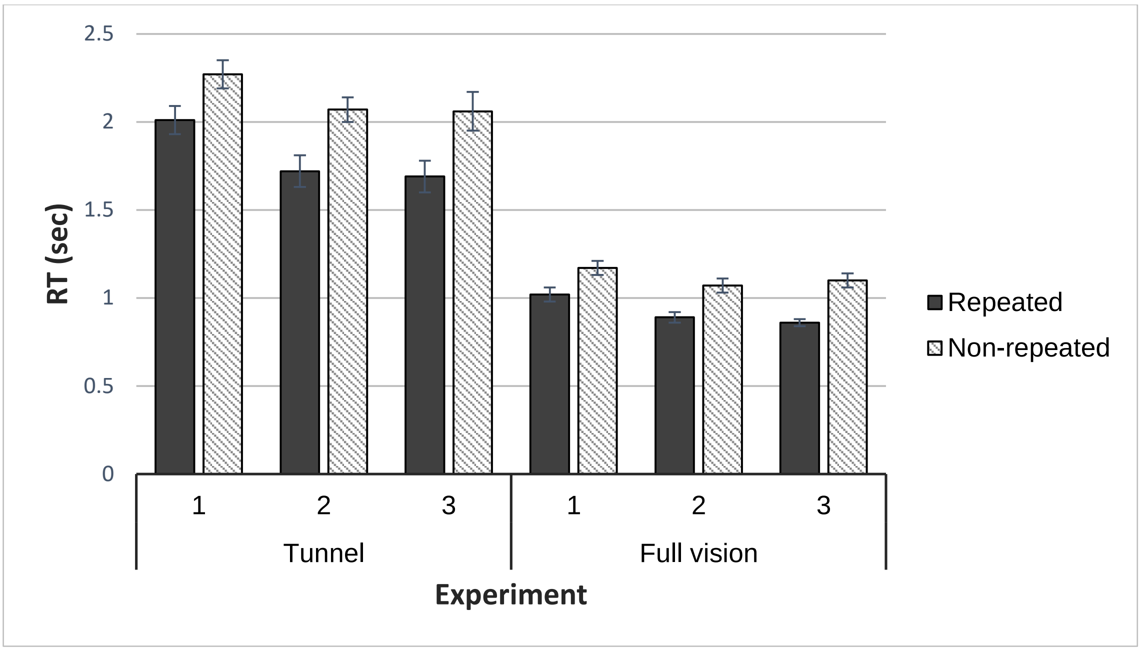


Figure 9. Mean of the reaction times (RTs) and standard errors for the repeated and non-repeated condition as a function of epoch (1-3), in tunnel vision group and full vision group.

2.2.2. Recognition test.

To be considered correct, the position selected by the subject (i.e. x,y coordinates given by pressing the mouse button) had to fall within the segment where the cup was originally presented in the main experiment.

The average recognition accuracy in the tunnel vision group was 61.30% (SD=18.74%). To test whether the participants were able to recall the position of the target, we compared the accuracy obtained in the sample with the chance level value of 16%, since in each trial the probability to select the right position was 1/6. Participants' accuracy was significantly above chance ($t(19)=10.81$, $p<0.001$), indicating that they recalled the exact target location in the repeated displays.

In the full vision group the accuracy was 71.35% (SD=21.12%). This, too, was significantly above chance ($t(19)=11.72$; $p<0.001$).

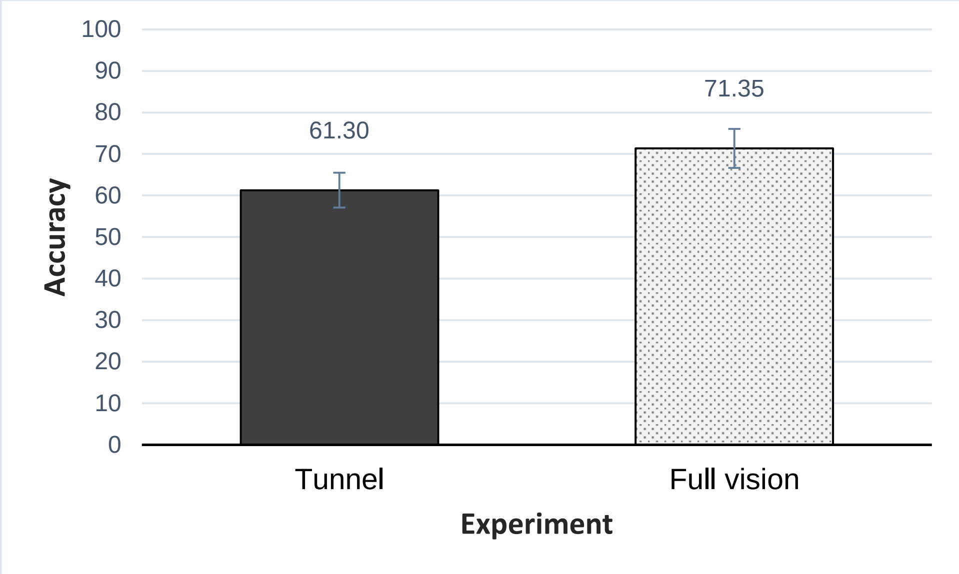


Figure 10. Mean of the accuracy (with standard errors) in the recognition test respectively in the tunnel vision group and in the control group.

Furthermore, the accuracy in the two groups was not significantly different ($t(37)=1.59$, $p=0.119$), that is, tunnel vision did not significantly influence the quality of the memory trace (Figure 10).

2.3. DISCUSSION

The aim of the present study was to investigate whether contextual cueing effects in real-world scenes can occur when visual search is performed with a gaze-contingent window (McConkie & Rayner, 1975; Reder, 1973). Over the years, two contradictory views have emerged regarding the type of spatial information that guides visual search in familiar environments the most. On one hand, there are multiple pieces of evidence that local spatial cues lead to strong contextual cueing (Brandy & Chun, 2007; Olson & Chun, 2002; Jiang & Wagner, 2004), whereas on the other hand, there is growing evidence that global context information plays a key role in contextual cueing (Brockmole et al., 2006; Rosenbaum & Jiang, 2013; Zang et al., 2015). In general, the concept of “local” and “global” can be difficult to define, since what is considered as local in one circumstance, can be seen as global within another (Brooks, Rasmussen & Hollingworth, 2010). For instance, in natural scenes, a table can be considered local context with respect to the scenario kitchen, while the same table can constitute the global context for the elements on the table itself. In order to attempt to solve this issue, in the present study the local information is clearly defined as the elements neighbouring the target. Compared to previous studies (Brady & Chun, 2007; Brockmole & Henderson, 2006), where both the local and the global information were continuously available during the visual search, in the present investigation, thanks to the gaze-contingent window, it was possible to selectively examine the role of the foveal information in contextual cueing.

Two groups of students were tested in a visual search task. Half of the participants performed the task with a gaze-contingent window that allowed them to see just the part of the visual field around the fixation, excluding the peripheral information. The other group performed the search without any viewing restriction. Therefore, the critical manipulation across groups was the capability of using the entire visual field, or, on the contrary, having access only to the local content of the display. The participants were asked to explore twelve 3D-rendered images of indoor rooms and search for a specific object in the room, that constituted the target of the search. The results of the experiment showed that when the scenes were repeated, that is, the position of the target was constant across repetitions, the response time was shorter than when the target was in a new position at each presentation. This indicates that contextual learning took place. Critically, both groups showed a benefit over time from the repetition of the displays, independently of the viewing condition. The effect arises from the learnt association between scenes and target position, that facilitated the visual search of the relevant target in the following expositions of the display. In other words, the memory representation interacted with visual attention to enable efficient visual search. This is supported by the subsequent memory task. The subjects of both groups proved to be aware of the repetitions, indeed they were able to identify the position of the target with significantly high accuracy, underlining the explicit nature of the memory mechanism behind contextual cueing. Also, the accuracy in the retrieval was not significantly different in the two viewing condition.

As previously reported (Brockmole & Henderson, 2006; Chun & Jiang, 1998) the repetition of spatial configurations can facilitate the finding of the target in both naturalistic and artificial displays. The present results are also consistent with previous investigations that suggested that attention is guided by the local information in proximity to the target (Brady & Chun, 2007; Olson & Chun, 2002). Since the window leaves open just the circumscribed portion of the visual field around the fovea, it is reasonable to affirm that just the local information around the target is processed by observers during the search. The results reported here are in line with other studies, investigating contextual cueing in letter arrays displays. Likewise, Brandy and Chun (2007) found that contextual cueing was based on the pairwise association between the target location and the distractors around it. A significant effect was indeed found when only the context near the target was processed. Maintaining constant a limited and highly localized portion of the display was sufficient for speeding up the search performance in repeated configurations (Olson & Chun, 2002).

Contextual cuing is considered a form of spatial statistical learning (Goujon, Didierjean, & Thorpe, 2015; Zellin, von Mühlenen, Mülle,r & Conci, 2013). Observers can reduce the cognitive complexity of the visual world by encoding and maintaining in memory the regularities they repeatedly encounter in the environment, and in turn adopt efficient behaviour. In order to determine the statistical regularities, the observers identify the statistical relationships between the configuration of the target location and the elements in its proximity

(Chun & Turk-Browne, 2008). For instance, in entering a bedroom, a person would probably expect the pillow on the bed, since a connection between the two elements was established over many time of encountering the two elements in association. Therefore, observers learn visual pattern where stimuli co-occur frequently (Baker, Olson, & Behrmann, 2004; Fiser & Aslin, 2002). Attention is necessary to develop statistical learning. Robust contextual cueing was indeed observed only when the repeated information was selectively attended, in a mechanism that guides attention toward relevant information (Jiang & Chun, 2001). Visual attention archives the prioritized processing of certain items of the visual field, usually considered relevant to the task, so that only attended objects can be discriminated and recognized. More precisely, whenever visual attention is allocated to a specific region of the visual field, the objects in it are perceptually processed and recognized, thereby stored in visual long-term memory (Deubel & Schneider, 1996). Objects that are more often fixated and for a longer period of time are typically remembered better (Hollingworth & Henderson, 2002). Ultimately, memory for spatial properties of a scene interacts dynamically with the perceptual processing of the items during visual search, guiding the attention to the target location and allowing the visual system to process that part of the visual display in high resolution (Hollingworth, 2006). All together, these considerations lead to the conclusion that attention is spatially focused around the target, encoding the local area around it.

Processing just the local information instead of the entire global context can be ecologically efficient especially in real-world scenes. In daily life, it seems more important to pay attention to the local context neighbouring the target, than to the large scene, because objects may covary more frequently with items in their local context, compared to items in the long-range context (Brady & Chun, 2007). Furthermore, statistical learning is an intensive process and it can be demanding applying it to the entire visual field efficiently, so in terms of behaviour it is more convenient to learn the contingencies restricted to the local region that brought a successful outcome (i.e. finding the target). Additional support of this notion is given by experiments that investigated the lack of foveal information in contextual cueing paradigm. It has been shown that the contextual cueing effect never developed for repeated displays with simulated central scotoma (Geringswald et al., 2012) or in patients suffering from age-related macular degeneration (Geringswald et al., 2013). On the contrary, other research suggested that peripheral information is necessary to efficiently retain the learnt spatial constraints, in order to facilitate visual search, because contextual cueing suffers under limited foveal viewing condition (Zang et al., 2015). A possible explanation is offered by Brockmole and colleagues (2006) that investigated contextual cueing in real-world scenes. The conclusion they arrived to is that global context is necessary to develop contextual cueing because observers have a strong bias in favor of associating the target to the global scene context, since it may contain more rich information.

It is noteworthy that the influence of foveal/local information and peripheral-global information was investigated differently in the studies mentioned above compared to the current one. In both letter arrays (Brady & Chun, 2007; Olson & Chun, 2002) and real-world scenes (Brockmole et al., 2006; Brooks et al., 2010) the entire display was in fact constantly available during the search, while here the gaze-contingent selectively excludes the foveal/local information from the peripheral/global information. With this investigation I hoped to fill the open gap regarding how spatial processing of foveal/local context may influence contextual cueing in real-world scenes. An interesting follow up would be to study how contextual learning may be influenced by encoding only the global context in real-world scenes, that is the peripheral visual field, by using a simulated scotoma.

In conclusion, the current findings brought evidence that context regularities can be learnt with a gaze-contingent window that leaves uncovered just the foveal portion of the visual field. The results show that scene-target location association was efficiently learnt in the tunnel viewing conditions as well as in the free viewing condition, so that finding the target in repeated configurations was faster compared to non repeated. Moreover, the memory that guides attention to the relevant part of the display is encoded explicitly, so that participants can access to that information in order to guide visual attention to the target in an efficient way.

3

CONTEXTUAL CUEING IN REAL-WORLD SCENES WITH SIMULATED SCOTOMA

The visual world is a rich environment that contains a large amount of visual stimuli, making it impossible to process everything with the same accuracy. To solve its complexity, visual attention developed to notice and remember the regularities in the environment, based on the knowledge acquired with experience (Henderson et al., 1999; Wang et al., 1994). The activation of a schema allows observers to predict what objects tend to co-occur with each other. Visual attention helps to focus on a specific subset of elements and facilitates the processing of relevant information over irrelevant (Chun & Wolfe, 2001). When a target possesses salient bottom-up features such as color, size, orientation or motion, it can draw attention to itself easily (Wolfe,1994; Yantis & Hillstrom, 1994). However, in natural scenes these types of cues are too numerous to guide attention, therefore they are often ignored by the observers (Simon & Levin, 1997). In complex scenes an important cue for attention is then the context, since it can be used to predict the configuration of the elements in it (Chun, 2000). For example, in the bedroom one typically finds a bed, wardrobe, dresser and bedside tables: each item cues with the presence of the others.

The context in rich, structured environments facilitates the processing of predictable configurations over random structures and guides visual attention to the task relevant elements. This effect is known as contextual cueing and proves that repeated presentations of specific arrangement of target and distractors lead to a more efficient target localization (Chun & Jiang, 1998). The classic contextual cueing paradigm involves a visual search task in which a rotated T is present among rotated Ls, the participants are asked to locate the target as quick as possible. Critically, half of the displays are repeated during the experiment, while the other half are new. Over repetitions, the participants became faster in locating the target in the repeated displays compared to searching for the target in new configurations. The effect was ascribed to a form of implicit learning, as it occurred without the participants being aware of the repetitions (Chun, 2000; Chun & Jiang, 2003; Brady & Chun, 2007).

Although very efficient, the letter arrays used by Chun and Jiang as stimuli, lacked of the realism that characterize natural scenes (Chun, 2003). Real-world environments have the benefit of being a relatively stable collection of objects that covary in a predictable way (Henderson & Hollingworth, 1999). Brockmole and Henderson (2006) used the constant relationships among objects in a room to, for the first time, apply contextual cueing to real-world scenes. Their study aimed to examine how regularities in the real-world environment can guide visual attention to relevant targets. To this end, their participants searched for an arbitrary letter inset into photographs of indoor rooms. While search time for new configurations was constant during the experiment, searching the target in repeated configurations was faster over repetitions.

Unlike previous results, the memory for scene-target location was explicit. When asked to recognize the repeated displays, the participants performed above chance (97% hit). Furthermore, the participants showed high accuracy in indicating the exact position of the target. The semantic constraint inherent in realistic scenes appeared to facilitate explicit learning and retrieval of target-context configurations. Therefore, the attention was guided to the target location in a more efficient way. The scene-target location association was learnt up to five time faster and led to a search time advantage twenty times greater compared to contextual cueing in letter arrays (Brockmole & Henderson, 2006). All these studies demonstrate that attention is facilitated in familiar environments and provides a benefit in terms of efficiency in searching for a target in repeated displays. Context guides eye movements so that observers direct their attention to the most important regions of the scenes with high-resolution foveal vision (Yarbus, 1967). When exploring a scene, observers make eye movements in order to fixate objects that were judged informative in the scene (Loftus & Mackworth, 1978; Henderson, 2003). An interesting question is what happens to contextual cueing when the foveation is compromised.

Geringswald et al. (2012) tested whether contextual cueing with symbolic displays can be affected by the presence of a gaze-contingent display that excluded foveal vision. The results indicated that under impaired viewing conditions, the participants did not show any benefit in searching for a target in repeated configurations. Foveal vision loss forced participants to explore the scene relying only on the information provided by peripheral vision, which is characterized by lower resolution and therefore, reduced visual acuity and contrast sensitivity (Westheimer, 1960; Campbell & Green, 1965). Moreover, the central scotoma interfered with eye movements, making the search more difficult (Cornelissen, Bruin, & Kooijman, 2005). In addition, the impaired viewing condition may have affected the attention and possibly the capability of creating an association between target location and display. Indeed, eye movements are closely related to attention (Deubel & Scheider, 1996; Hoffman & Subramanian, 1995), therefore the impossibility of fixating with the high resolution of the fovea might have compromised the localization of the target . As a consequence, this could have led to a lack of visual memory, since fixating an object is linked to its recall (Hollingworth, 2006).

On the other hand, several studies investigating contextual cueing in natural scenes demonstrated the importance of global context in visual search (Brockmole, Castelhano, & Henderson, 2006; Rosenbaum & Jiang, 2013). Brockmole and colleagues asked participants to search for a target letter embedded in 3D-rendered illustrations of real-world environments (i.e. interior rooms). Critically, the local (i.e. the group of objects in close proximity to the target) and the global context (i.e. the scene in background) were dissociated, so that half of the participants saw the altered local context while preserving the global context, whereas the other half saw the changed global context while maintaining the local context. Contextual cueing was preserved when the local context was modified, however, changes in the global context completely eliminated the benefit of the repeated configurations. Additionally, the scene-target association was learnt faster and better when the global context was repeated across trials. In a similar vein, Rosenbaum and Jiang (2013), proved that participants were more likely to encode the target location with the global scene instead of the local context. Contextual facilitation was preserved only when the background scene was predictive of the target location.

In the present investigation, I tested whether the global context provided by the real-world scenes, removes the need to attend the items with central vision. The aim of the experiment is, therefore, to ascertain whether the information coming from the peripheral vision can efficiently improve visual search in repeated displays. For this purpose, I compared the performance of two groups of students, searching for a target in real-world environments, critically, one group performed the task with a gaze-contingent scotoma (Rayner & Bertera, 1979) that excluded foveal vision, while the other group searched for the target without any viewing restriction.

Since global information in real-world scenes proved to be an important factor for learning contextual cues, it is possible to hypothesize that peripheral information can be sufficient to generate the contextual cueing effect in presence of foveal vision loss.

3.1. METHOD

3.1.1. Participants

Scotoma group. Twenty (10 males, 10 females; average age 25.55 , sd= 3.20) undergraduate students of the Otto-von-Guericke University of Magdeburg participated in the study.

Full vision group. Twenty (1 male, 19 females; average age 20.35 , sd= 1.93) undergraduate students of the Otto-von-Guericke University of Magdeburg took part in the experiment as the control group.

All participants reported normal or corrected-to-normal visual acuity and were not aware of the purposes of the study. Written informed consent was provided before the beginning of the experiment, which was approved by the Ethics Committee of the University of Magdeburg. Participants chose to be compensated with course credits or with a remuneration of 6 euros/hour.

3.1.2. Apparatus

The stimuli were presented and the responses recorded with PsychoPy version 1.82 and a PC running Debian Linux. The monitor connected to the PC was a LCD BenQ XL2410T full color 24 inch HD, it was 521 mm (1920 pixels) wide and 293 mm (1080 pixels) high, with a refresh rate of 120 Hz. The stimuli were presented at a distance of 80 cm in a quiet and dimly lit room.

3.1.3. Stimuli

Stimuli consisted of 3D rendered illustrations of twelve indoor rooms. All the images were created with an open source interior design software (Sweet Home 3D, version 4.6, eTeks Paris, France. Available at http://www.sweethome3d.com. See Appendix) and represented the following twelve different rooms: a bathroom, a bedroom, a cinema room, a children's room, a garage, a recreation room, a kitchen, a library, a living room, a music room, an office and a study. Each type of room was presented in a singular exemplar. The rooms were created balancing the colors of the environment and the number of objects in the room, in order to avoid any pop-out effect of the target. The scenes were displayed at a resolution of 1200x900 pixels, subtending a visual angle of 22.83x17.12 degrees. Each scene included a yellow cup, which constituted the target in the visual search task. The cup's position was equally allocated to one of the six equal-sized rectangular segments in which the scene was divided, three (left, middle, right) above respectively below the horizontal midline. The cup was always presented in a meaningful position (e.g. cup on a surface and not floating in the air) and clearly visible.

3.1.4. Design and procedure

The experiment was divided into two parts, separated by a short break.

Visual search. During the course of the first part of the experiment, participants viewed 72 stimuli, divided into six blocks of twelve trials each. Each block contained six scenes in which the position of the target was repeated throughout the experiment (i.e. the repeated condition), in the remaining six scenes the target was presented in a different segment at each block (i.e. the non-repeated condition). Each block randomly intermixed the repeated and the non-repeated trials. For each participant, the sequence of the stimuli was randomly assigned and the rooms that constituted the repeated displays were randomly drawn from the pool of all the rooms. The subjects were not instructed that in some scenes the target's position was repeated during the experiment. Participants were asked to search for the yellow cup in each scene and to indicate the side of the handle by pressing the left or the right button of the mouse as quickly as possible. The side of the handle varied randomly, both in novel and repeated displays, so that no association of scene and response could be learnt across the repetitions. At the beginning of each trial, a fixation cross was presented at the center of the screen for 1000 ms in black on a gray background. Participants were instructed to fixate the cross. After the fixation cross, a blank gray screen was presented for 500 ms and after that, the search display was shown and remained visible until the response. Upon identifying the target, the subjects had to press the left button of the mouse when the cup had the handle on the left side or the right button of the mouse when the cup had the handle on the right side. After the response, a feedback sound was provided (i.e. a high pitch tone for the correct response and a low pitch tone for the wrong response).

In the experimental group (i.e. the scotoma group) a gaze-contingent scotoma was simulated (Rayner & Bertera, 1979). The gaze-contingent display was created by displaying a solid gray circle with a radius of 5° of visual angle, covering the foveal and parafoveal region of the retina (Beard and Ahumada, 1999; Coletta and Williams, 1987; Larson and Loschky, 2009; Rayner et al. 1981). Throughout each trial, the position of the participant's left eye was recorded at the rate of 1000 Hz by an Eye-Link 1000 eye-tracker system (SR Research, Inc.), and according with the coordinates (x,y) provided by the eye-tracker device the gaze-contingent display was drawn. When the gaze coordinates were unavailable due to eye blinks or signal losses, the screen was kept completely blank. Before each block a 9-point gaze calibration was performed, followed by a 9-point gaze validation at the beginning of each trial. The calibration and the validation procedure were repeated in case the error was bigger than 1° on average or bigger than 1.5° for the worse point. The scotoma simulation was not presented neither during the calibration nor the validation procedure (Figure 11).

Recognition task. In the second part of the experiment, participants completed a memory task, in order to test whether the memory for scene-target position was explicitly encoded. After a short break, the subjects performed a recognition task that consisted of the six real-world 3D rendered rooms presented in the repeated condition. The scenes were now rendered without the search target (i.e. the yellow cup) and participants were asked to move the mouse cursor and click on the target location. At the beginning of each trial, a fixation cross was presented at the center of the gray background for 1000 ms and then the scene was displayed until the response. No feedback was provided.

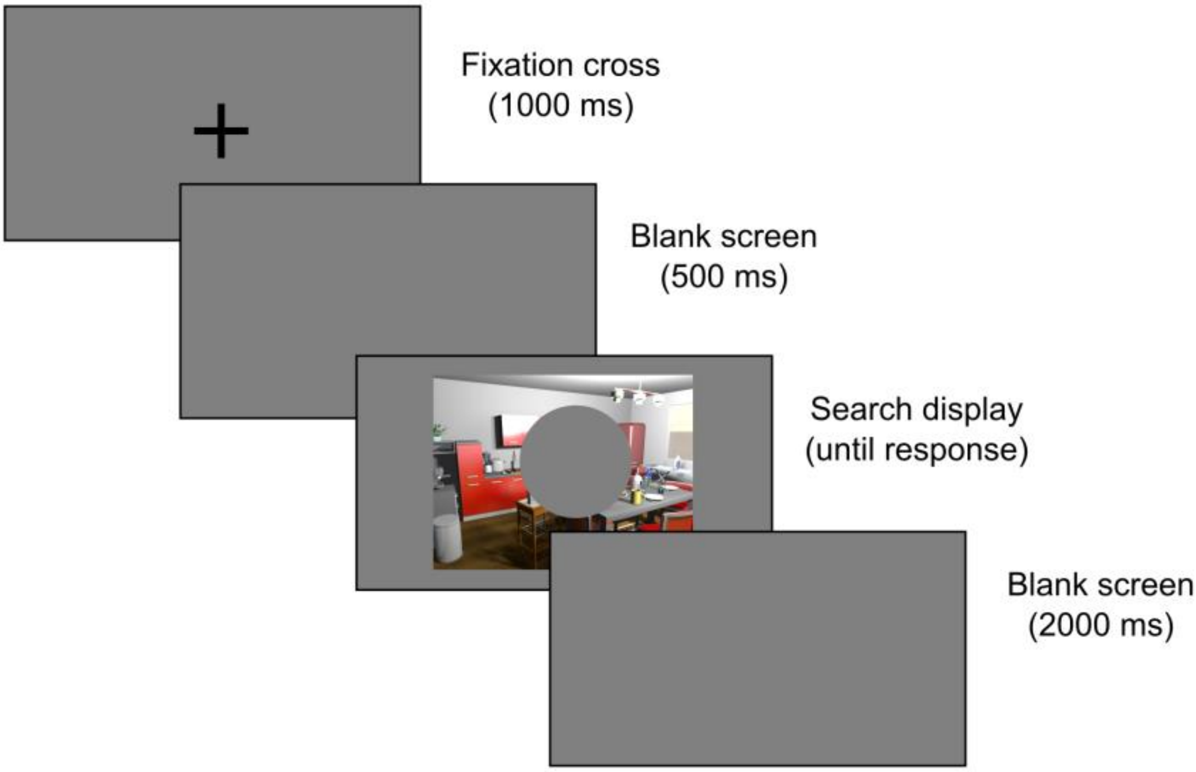


Figure 11. Example of a visual search trial with the gaze-contingent scotoma.

Scotoma validation task.

To validate the simulated scotoma, each subject in the scotoma group performed a search-discrimination task (Geringswald, Baumgartner, & Pollmann, 2013). The experiment consisted of 4 blocks with 32 stimuli each; in the first and third block the participants searched for the target without the central scotoma simulation, in the second and fourth block participants looked for the target with the simulated central scotoma. The gaze-contingent display used was the same as in the main experiment (i.e a solid gray circle with a radius of 5° visual angle).

At the beginning of each trial, a fixation cross was presented at the center of the screen for 1000 ms, followed by the search display, which consisted in a Landolt C (i.e. a circle with a gap at the left, top, right or bottom) used as the target, among seven same-sized rings used as

distractors. These stimuli were placed on an invisible circle with a radius of 10° visual angle. The stimulus remained on the screen for 5000 ms or until the response. The participants were asked to find the Landolt C and to indicate the side of the gap (e.g. left, top, right or bottom) by pressing the corresponding arrow key in the keyboard. After the response a feedback sound (a high pitch tone for correct responses, a low pitch tone for incorrect responses) was provided. The position of the Landolt C and the direction of the gap were randomized among subjects.
The size of the C was chosen so that the opening (extending 0.067°) should not be discernible with peripheral vision in the presence of the central scotoma (see Geringswald et al., 2013 for details).

3.2. RESULTS

3.2.1. Visual search

Accuracy. In both groups the accuracy during the search task was high. On average, in the experimental group the accuracy was 93.82%, whereas in the normal vision group it was 99.51% ($t(21)=5.12$, $p<0.001$). In the scotoma vision group, one participant was excluded from further analysis because of the excessive errors (i.e. 23.6%). This subject was replaced with a new one, who performed the exact same sequence.

Response time. Only the correct responses were considered in the further analysis and the outliers were excluded when greater than two standard deviations from the mean computed on a subject-by-subject basis. For statistical analysis a log transformation was performed on the data and a three-way mixed design ANOVA was run to calculate the contextual cueing effect, with Group (scotoma vision vs full vision) as between-subjects factor, Configuration (Repeated, Non repeated) and Epoch (1-3) as within-subjects factors. With a total of six blocks, each epoch consisted of two blocks (Figure 12). A strong effect of Group emerged ($F(1,38)=$

180.7, $p < 0.001$), reflecting the threefold longer search time in the group of participants forced to deal with the scotoma simulation. That is, response time was faster in the full vision group (1019 ms) compared to the scotoma group (3505 ms). Moreover, the significant main effect of Epoch ($F(2,76)=14.55$, $p < 0.001$) indicates that both groups showed a general improvement over time and the significant main effect of Configuration ($F(1,38)= 24.55$, $p < 0.001$) suggests that the search time was faster in the repeated displays. No interaction was significant ($Fs<2.73$, $ps<0.071$). Critically for the contextual cueing paradigm, the interaction Epoch x Configuration did not reach the significance ($F(2,76)=2.73$, $p= 0.071$). Even though the interaction was not significant, it seemed reasonable to clarify the trend by running an additional one tailed t-test, in order to assure that the search difficulty was initially equal in both conditions. While there was no significant difference in the first epoch ($t(76)= 1.35$, $p=0.09$), the repeated condition showed a significant improvement in the last epoch ($t(76)=2.52$, $p= 0.007$), compared to the non repeated condition, thus reflecting contextual cueing during the experiment. In the non repeated condition little change in the search efficiency was found across epochs, from 3296 msec in the first epoch to 2895 ms required in the last epoch. In contrast, the repeated condition revealed a substantial decrease in search time (i.e. 1.022 ms), the target was located in 2381 msec at the beginning of the experiment, whereas only in 1359 ms at the end. Furthermore, there was no reliable interaction between Group and other factors, suggesting the two groups produced a not significant different pattern in contextual cueing effect ($Fs>2.83$, $ps>0.101$).

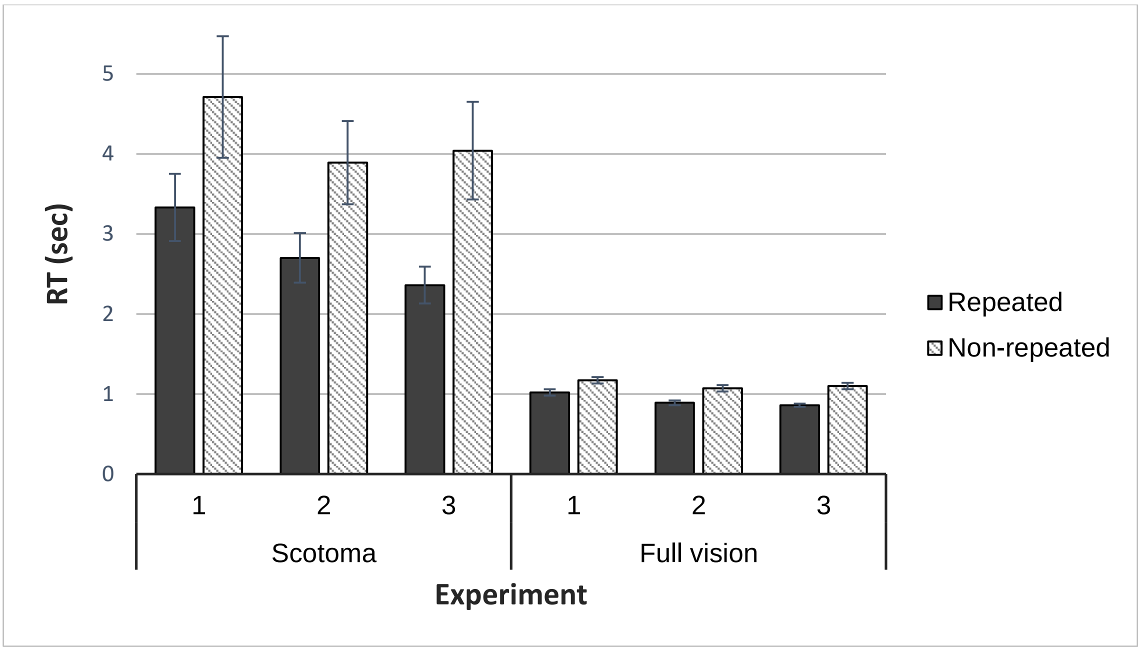


Figure 12. Mean of the reaction times (RTs) and standard errors for the repeated and non-repeated condition as a function of epoch (1-3), in scotoma group and full vision group.

3.2.2. Recognition test

To be considered correct, the position selected by the subject (i.e. x,y coordinates given by pressing the mouse button) had to fall within the segment where the cup was originally presented in the main experiment.

The average recognition accuracy in the scotoma group was 53.10% (SD= 24.65%). To test whether the participants were able to recall the position of the target, we compared the accuracy obtained in the sample with the chance level value of 16%, since in each trial the probability to select the right position was 1/6. Participants' accuracy was significantly above chance ($t(19)=6.73$; $p<0.001$), indicating that they recalled the exact target location in the repeated displays.

In the full vision group the accuracy was 71.35% (SD= 21.12%). This, too, was significantly above chance ($t(19)=11.72$; $p<0.001$).

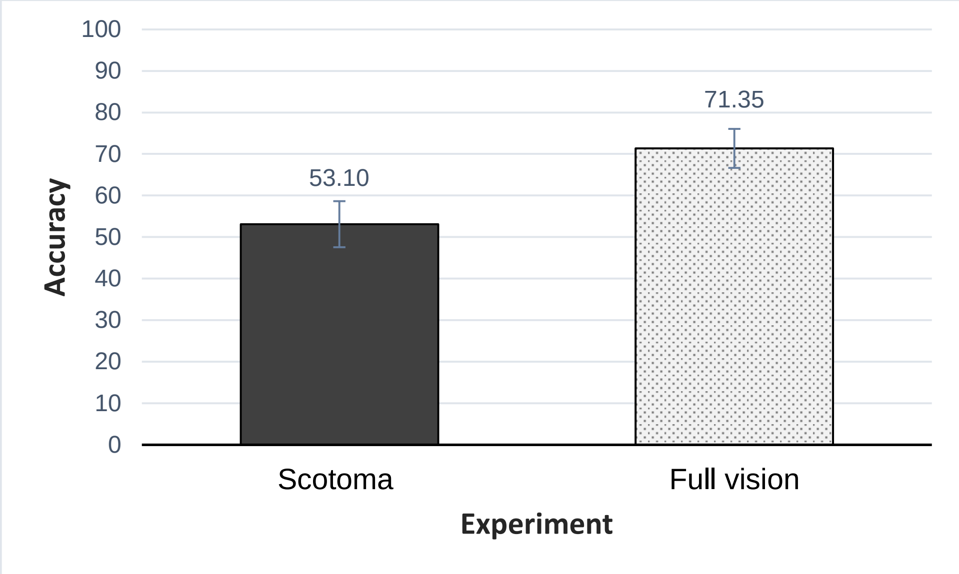


Figure 13. Mean of the accuracy and standard errors in the memory task for the scotoma group and the control group.

Furthermore, the accuracy in the two groups was significantly different ($t(37)=2.51$, $p=0.016$), indicating that the scotoma influenced the quality of the memory trace (Figure 13).

Scotoma validation task

In order to test the validity of the gaze-contingent display used in the main experiment, we compared the response accuracy in the two viewing conditions. Without scotoma simulation, on average, the direction of the gap was identified correctly in 95% of the trials (i.e. 60.65 trials out of 64), whereas in the condition with the gaze-contingent display, only 31% correct responses (i.e. 19.95 trials out of 64) were recorded. Thus, the performance dropped significantly ($t(19)=20.78$, $p< 0.001$), reflecting an impairment in the stimulus discrimination when the vision was compromised by the scotoma simulation.

3.3. DISCUSSION

In the present study, I explored whether the loss of foveal vision affects people’s ability to extract and remember consistent spatial connections in real-world scenes. To this end, two groups of students were tested in a visual search task. One group executed the task with a gaze-contingent scotoma that prevented them from using the central part of the visual field during the search, therefore receiving information only from the periphery of the visual field. The other group completed the task without any visual restriction. For this purpose, a visual search task was designed in which participants had to explore twelve 3D-rendered images of naturalistic environments and search for a pre-specified target in the rooms. The important manipulation across groups was the capability of using the central part of the visual field, as we normally do, or on the contrary, being forced to rely only on the information coming from peripheral visual field. In order to determine on which memory mechanism contextual cueing relies upon, participants were afterwards tested on a recognition task, although they were not previously instructed to memorize the stimuli. To test the validity of the gaze-contingent scotoma in blocking the foveal processing, each subject in the scotoma group performed an additional task. This secondary task had the purpose of controlling whether the scotoma effectively affected the performance of the subjects when the high resolution of the fovea was necessary to complete the task successfully (for details see Geringswald et al., 2013). Therefore, the accuracy in discriminating a Landolt C among other rings was measured with and without scotoma. Accuracy dropped from almost perfect discrimination in the unimpaired viewing condition to a compromised discrimination with the scotoma simulation. Since the performance was closely related to the capability of using the high acuity provided by the foveal part of the visual field, the scotoma effectively interfered with the foveal processing.

The results of the visual search experiment showed that overall, all participants improved their performance during the experiment, becoming faster in finding the target in the scenes. When configurations were repeated, however, response time was shorter. Moreover, the difference in the response time between repeated and non-repeated trials was larger at the end of the experiment compared with the beginning, indicating that the learning of the contingencies between target and scene has taken place, therefore revealing the contextual cueing effect. Although the contextual cueing effect was affected by the presence of the gaze-contingent scotoma, a significant improvement over time for the repeated condition was found. These findings suggest therefore that the foveal vision loss does not prevent contextual cueing to occur in real-world, nameable scenes. When asked to recall and specify the position of the target in the scenes, participants performed above chance level in both groups, indicating the ability of recalling the correct position of the target in the repeated scenes. Therefore, the memory for target position-scene covariation was explicitly encoded. Although the scotoma group recalled the association scene-target position explicitly, it did not perform as well as the participants that explored the scene with full vision. The accuracy in the scotoma group was indeed significantly lower, indicating that participants struggled in the retrieval of the target position. This result suggests that the memory trace was weakened by the loss of central vision. Since the high acuity resolution of the fovea was compromised by the gaze-contingent scotoma, the participants might have experienced some difficulty in discriminating the fine details in the scene, therefore this impacted the quality of the retrieval. The inefficient retrieval in the scotoma group compared to the controls may have been a possible reason why the interaction Epoch x Configuration barely did not reach the significance. The reduced memory for the target position may have compromised the stimulus-guided search, which resulted in a diminished search facilitation compared to the unimpaired viewing condition.

The results presented here support previous research that investigated contextual cueing in real-world scenes. Brockmole & Henderson (2006) demonstrated that a repeated exposure to real-world scenes led to decreased search time for targets located consistently in the same position, suggesting that subjects learnt the association scene-target location over repetitions. Compared to visual search in arrays of letters (Chun & Jiang, 1998), the cueing in real-world scenes occurred faster and was explicitly encoded in memory, so that the subjects were aware of the repetitions and recalled the position of the target with significant high accuracy. The memory for the repeated visual context indeed, predicted the location of the target and ultimately guided the attention efficiently to the region that was relevant for the task. As a result, the search time was shorter for the repeated configurations compared to the new ones. Compared to symbolic arrays, the identity of real-world environments provides semantic constraints that help the observer to generate expectations about the objects and the relationships among them, as scenes are collections of elements that covary in a predictable way (Henderson & Hollingworth, 1999). Critically, Brockmole, Castelhano and Henderson (2006) showed that learning benefit was, indeed, preserved when the global scene information was kept intact, however, contextual cueing was completely eliminated when only the objects surrounding the target remained unchanged. This proves that observers, in real-world scenes, associate the target location with the scene identity more than with the elements neighboring the target. Global context better enable the scene identity because it supplies a rich amount of information about the elements in the scenes. Thus, global context allows observers to activate a schema, which facilitates the analysis of the objects arranged in the scene in a consistent manner. Furthermore, global context provides a recovery of knowledge from the past experiences faster than the small subset of elements placed around the target, allowing a quicker transfer of learning (Brockmole et al., 2006). A key component of contextual cueing is, indeed, the retrieval of previous memories. The visual system matches the incoming display with the previous memory trace: the easier the match, the faster the attention is guided to relevant spots (Song & Jiang, 2005). Opposite results were found in symbolic displays (i.e. T among Ls), where observed appeared to rely preferentially on local information. Destroying the elements near the target compromised contextual cueing, whereas, global configuration did not bring any significant benefit (Brady & Chun, 2007; Olson & Chun, 2002). These results suggest that the importance of the spatial proximity is specific for letter arrays and does not hold for natural scenes. The types of stimuli used in the experiment appear to influence the way the context is encoded.

Contextual cueing has been studied extensively with symbolic displays and real-world scenes in free viewing conditions, but the role of pure peripheral/global information in contextual learning in realistic scenes is still unclear. To account for the lack of knowledge in this particular aspect, I examined whether contextual cueing was developed with a central scotoma simulation, that blocked central vision. In previous experiments using symbolic displays (i.e. letter arrays), it has been found that central vision loss severely interfered with the learning of spatial contextual cueing. The scotoma prevented participants to use the learnt configurations to improve the efficiency of visual search, as a result, the facilitation in repeated displays was eliminated. The lack of contextual cueing was observed with both gaze-contingent scotoma (Geringswald et al., 2013; Geringswald & Pollmann, 2015) and natural scotoma (macular degeneration; Geringswald et al., 2012). Geringswald et al. (2012) tested a group of normal-sighted students searching for the target with a gaze-contingent display that masked the central part of the visual field. The presence of the scotoma was found to impair the contextual cueing in repeated displays. Thus, the authors concluded that the gaze-contingent scotoma prevented the participants from using the learnt configurations for an efficient visual search. The inefficient top-down controlled search, given by the uncommon viewing condition, may have made the match between the previous learnt displays and the current one more difficult, reducing the search facilitation in the repeated configurations. In order to compensate for the highly artificial situation of searching with a simulated scotoma, the same visual search task was presented to patients with age-related macular degeneration (AMD). In line with the previous finding, patients with naturally impaired foveal vision failed to show any search facilitation in repeated displays, while contextual cueing was preserved in the age-matched control group (Geringswald et al., 2013). The authors concluded that the loss of foveal vision impaired search facilitation in familiar environments. In contrast, the idea that contextual cueing could take place in the periphery of the visual field was investigated by van Asselen and Castelo-Branco (2009). In their experiment, robust contextual cueing effects were found, despite the fact the participants were instructed to fixate a fixation cross at the centre of the display, that is, to use only peripheral vision to conduct the search. These findings brought evidence that detailed foveal vision was not necessary to develop contextual learning. The current study supports what was previously reported by van Asselen and Castelo-Branco (2009), that is, we do not need detailed foveal perception to detect a target that appears in repeated configurations more quickly. In addition, the results of the current research show that in real-world scenes, visual search benefits from the exposure to repeated configurations also in the presence of a gaze-contingent scotoma.

Visual search performed relying exclusively on peripheral vision has multiple disadvantages because the peripheral part of the visual field processes the scene with low spatial resolution, therefore the fine details cannot be detected. In addition, searching with only peripheral vision is an uncommon practice. The gaze-contingent scotoma may have affected the way the participants explored the scene, which possibly have had repercussion on the capability of creating the association scene-target location and, therefore, its memorisation (Yarbus, 1967). A possible explanation for the preserved contextual cueing found in the experiment, lies in the nature of the stimuli. The need for top-down controlled eye movement planning in search with central scotoma may compete for visual working memory resources (Annac, Manginelli, Pollmann, Shi, Müller, & Geyer, 2013; Manginelli, Langer, Klose, & Pollmann, 2013), limiting the resources available for memory-driven search (Geringswald & Pollmann, 2015). This limitation appears less problematic in search in realistic displays, because the context provides expectations about the scene's layout and guides the subject's attention more efficiently to the target (Brockmole & Henderson, 2006). Consequently, in real world scenes the learning of repeated configurations interferes less with visual working memory (Vickery, Sussman, & Jiang, 2010).

Previous demonstrations of contextual cueing using symbolic displays have shown that observers were unable to discriminate the repeated arrays from the new ones, establishing that contextual cueing was based on implicit memory (Chun & Jiang, 1998; 2003; Brandy & Chun, 2007). However, in real-world scenes the association target position-scene was found to be explicitly encoded in memory, so that participants performed above chance when ask to discriminate the repeated displays from the new ones. Also, participants showed high accuracy in recalling the position of the target in repeated displays (Brockmole & Henderson, 2006). Observers are indeed able to recognize hundreds of pictures after a single exposure (Standing, 1973; Shepard, 1967). The awareness of the repetition in naturalistic stimuli may be explained by the fact that complex scenes are easier to discriminate from each other, since the variety of elements in it is greater compared to artificial arrays. In addition, the semantic categorization of the items contained in the scene is easy, for example, once observers are exposed to a picture of a bedroom, they will quickly categorize the scene as a bedroom and they will expect and remember a pillow on the bed.

The present investigation had the purpose of assessing whether peripheral/global information would be sufficient for developing contextual cueing in naturalistic scenes. To this end a gaze-contingent scotoma was simulated. Impaired central vision resembles a clinical condition called age-related macular degeneration (AMD), in which the patients progressively lose central vision. Future research may investigate whether AMD patients show a different results in contextual learning and if they have developed strategies to adjust search behaviour to compensate for the lack of central information.

In conclusion, this study uncovers the fact that in real-world scenes, peripheral vision can detect contextual information and the repeated configurations are used to guide attention and speed up visual search. The peripherally encoded information is explicitly retained in memory, however, the efficient retrieval of the learnt configurations is compromised by the central vision loss.

4

CONTEXTUAL CUEING IN PATIENTS SUFFERING FROM AGE-RELATED MACULAR DEGENERATION (AMD) USING REAL-WORLD SCENES

The environment we live in contains a large number of objects, and in order to compensate its complexity, the visual system notices and remembers the elements that covary together in a predictable way (Henderson & Hollingworth, 1999). These regularities allow the visual system to select the relevant information and recognize it across time, so as to reduce the serial scanning of the scene. When we enter a familiar environment, eye movements can be guided by memory of similar scenes. For example, when you think of your kitchen, you can explicitly tell where the refrigerator is and in which cupboard you will find a coffee mug. Similarly, when participants were asked to find a target in a realistic scene, search time was dramatically reduced if the target in the scenes was always placed at the same location (Brockmole & Henderson, 2006). This phenomenon has already been studied repeatedly in the past and it is known as "contextual cueing" (e.g. Chun & Jiang, 1998, 1999; Chun & Nakayama, 2000; Olson & Chun, 2001). In their work, Chung & Jiang (1998) demonstrated that the repeated exposure to the same arrangement of the target and distractors leads to a faster and more efficient visual search. The paradigm they used was an array of a rotated T (target) among several rotated Ls (distractors). Critically, across repetitions some random generated novel displays were alternated with a subset of repeated displays, that is, displays in which the target position was kept constant. The search time for finding the target decreased in a significant way in the repeated displays compared to the novel displays. The effect emerged without the participants being aware of the repetitions, indeed the subjects were not explicitly able to recognize the old displays from the new ones, nor could they identify the position that the target occupied in the display (Chun & Jiang, 2003). Spatial context learning was also investigated in real-world environments (i.e. indoor rooms; Brockmole & Henderson, 2006). Consistently with what found by Chung and Jiang (1998), participants proved to be sensitive to the relation between target position and scene, hence across repetitions search time required to find the target decreased in the repeated scenes. In real-world scenes, however, memory for the target position in the scene was explicit, that is participants recognized the repeated displays from those they saw once and recalled the position of the target with significant accuracy. The implicit nature of contextual cueing has been debated (Smyth & Shanks, 2008; Hout & Goldinger, 2010; Williams, Henderson, & Zacks, 2005; Brockmole, Castelhano, & Henderson, 2006), and now it is well known that, although contextual cueing in symbolic displays appears to be due to incidental learning and it occurs in the absence of explicit recognition (Chun & Jiang, 2003; Schacter, 1994), the contextual cueing in realistic scenes draws on explicit memory (Brockmole & Henderson, 2006; Brockmole et al., 2006).

In visual search foveal and peripheral vision play a complementary role. The peripheral vision locates a possible target within a scene and programs the eye movements, which allow the fovea, the region of the visual field with higher visual acuity, to scrutinize the target in its fine details (Bouma, 1970). The role of foveal vision during visual search tasks can be appreciated studying contextual cueing effects in patients suffering from age-related macular degeneration (AMD). Age-related macular degeneration (AMD) is an ocular disease that occurs mostly in the elderly and causes damages to the macula, the region in the retina with the densest concentration of cones, thus with the highest resolution power. AMD leads to the progressive loss of the central vision (Querques, Avellis, Querques, Bandello, & Souied, 2011; Klaver, Wolfs, Vingerling, Hofman, & de Jong, 1998; de Jong, 2006). Previous studies have shown that contextual cueing was severely reduced when participants were not able to rely on the information coming from central vision, due to a gaze-contingent display, that prevented them seeing the central part of the visual field (Geringswald, Baumgartner, & Pollmann, 2012) or because of a natural foveal vision loss (macular degeneration; Geringswald, Herbik, Hoffmann, & Pollmann, 2013). These findings suggested that the inability to orient the attention to the target, because of the lack of foveal vision, resulted in a reduced efficiency in guiding the search to the location of task-relevant information. Supporting the idea that perceiving a small portion of a display is not enough to create a strong visual memory (Zang, Jia, Müller, & Shi, 2015). Central vision loss seemed to interfere with the search facilitation because visual attention is closely related to the foveation of the target and its context, therefore the performance decreased since the eyes were not directly foveating the target (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995). Furthermore, learning in contextual cueing appears to be associated with the limited and highly localized area of the display near the target, so that the observers preferentially encode local context to guide attention to the relevant parts of the display (Brady & Chun, 2007).

On the other hand, several publications on vision-related quality of life in patients with AMD reported a high performance rate in more ecologic tasks that would require central vision, like face recognition and object recognition (Boucart, Dinon, Despretz, Desmettre, Hladiuk, & Oliva, 2008; Boucart, Despretz, Hladiuk, & Desmettre, 2008; Boucart, & Tran, 2011; Tran, Rambaud, Despretz, & Boucart, 2010). A recent study showed that in scene gist recognition tasks, the accuracy for scene categorization, although lower than in the age-matched controls, was significantly high in patients with AMD (around 80% hits), suggesting that peripheral vision is sufficient for recognizing a natural scene and categorizing it in terms of natural vs urban environment (Tran et al. 2010). In addition, Boucart and colleagues (2008) demonstrated that AMD patients showed an intact ability to rapidly categorize facial emotion, likely founding their decision mainly on the low spatial frequencies in the periphery, as they lack the perception of finer details. The studies mentioned above support the idea that a more naturalistic setting can help patients suffering from AMD to overcome the difficulties of exploring the environment without central vision. For this reason, it is possible to argue that patients showed a reduced contextual cueing because of the nature of the stimuli used in the experiment (Geringswald et al., 2013). Indeed, an array of letters is a really artificial configuration where the stimuli are not integrated in a meaningful environment and the lack of semantic relations can affect the strategies used in the search. The non-meaningful arrangement of the object might have prevented the patients from searching efficiently for the target, because the abstract stimuli did not allow them to transfer the exploration strategies they normally use in everyday life (Geringswald et al., 2013).

The aim of the present study was to investigate whether the additional semantic context provided by using real-world scenes would enable AMD patients to show contextual cueing effect, despite the central vision impairment. Also, contextual facilitation and its retrieval are investigated in comparison with a healthy age-matched control group.

4.1. METHOD

4.1.1. Participants

AMD Patients: We collected data from fifteen patients (8 males, 7 females; average age 77.6) diagnosed with age-related macular degeneration (AMD). Five patients were excluded due to clinical conditions (i.e. a patient was excluded because they were diagnosed with glaucoma other than the AMD; a patient was excluded because they could not perform the task monocularly, since the AMD was too severe) or technical problems (e.g. data loss during the experiment). Seven of the patients were recruited and diagnosed at the Ophthalmic Department of the University Clinic of Magdeburg, the other eight were recruited and diagnosed at the Augen Zentrum am Johannisplatz in Leipzig. Macular degeneration was determined based on an ophthalmic examination provided by the eye clinics. At the University Clinic Magdeburg, the visual field was tested with a black-on-white perimetry (Haag- Streit Octopus 101, dynamic strategy, size III, program dG2) and a microperimetry (NIDEX - MP1, 2-4 dB, Goldmann size III) to determine the locus of the fixation. At the Eye Clinic in Leipzig, the patients' central visual field was monitored by using the Amsler grid to detect the presence of metamorphopsia. All patients were affected by AMD in both eyes, but the progress of the pathology varied across subjects (For details see Table 1). Monocular testing was carried out on the more severely affected eye whenever possible (i.e. 7 patients), or in the less affected eye, when the subject could not perform the task with the other eye.

Control group. Twelve subjects (4 males, 8 females; average age 72.6, SD= 3.94) participated in the study as a normal-sighted (or corrected to normal) control group. Acuity was tested with the FrACT test (Bach M., The Freiburg Visual Acuity Test, version 3.9.8. Available at http://www.michaelbach.de/fract/index.html) prior to the experiment. For all the participants the decimal visual acuity was ≥ 0.84.

All participants were not informed about the purposes of the study. A written informed consent was provided before the beginning of the experiment, which was approved by the Ethics Committee of the University of Magdeburg. Participants received a remuneration of 20 Euros.

Table 1. AMD Patients information and clinical data.

Subject	Gender	Age	Type of AMD		Visus		Scotoma (°)		Fixation	
							horizontal	vertical	2°	4°
			RE	LE	RE	LE	RE	LE	RE	LE
S01	F	72	Dry	Dry	0.7	1.0	10, 7	4, 3	s (100%;100)	s (98%;99%)
S02	M	80	Dry	Wet	1.0	0.08	12, 13	/	i (31%; 71%)	/
S03	M	81	Dry	Dry	0.8	0.5/0.6	/	/	s (96%;100%)	i (61%;90%)
S04	M	74	Dry	Dry	0.04	0.5	3, 5	/	i (13%; 41%)	s (99%;100%)
S05	M	76	Dry	Dry	1.0	0.25	/	/	i (71%;100%)	i (61%; 97%)
S06	M	80	Dry	Dry	0.32	0.08	1, 1	/	i (40%; 87%)	/
S07	F	74	Dry	Wet	0.8	0.16	/	5, 7	s (100%;100)	i (44%; 95%)
S08	F	80	Dry	Dry	0.5	0.32				
S09	F	84	Dry	Dry	1.0	0.8				
S10	M	71	Dry	Dry	0,63	0,63				
S11	F	78	Dry	Wet	0.8	0.4				
S12	M	79	Dry	Dry	0,63	0,63				
S13	F	75	Dry	Dry	0.03	0.8				
S14	M	80	Dry	Dry	0,63	0,63				
S15	F	79	Dry	Dry	0.8	0.8				

Notes: M= male; F= female; RE= right eye; LE= left eye; s= stable; i= unstable.

4.1.2. Apparatus

The stimuli were presented and the responses recorded using PsychoPy version 1.82 and a PC running Debian Linux. The monitor connected to the PC was a LCD BenQ XL2410T full color 24 inch HD, it was 521 mm (1920 pixels) wide and 293 mm (1080 pixels) high, with a refresh rate of 120 Hz. The stimuli were presented at a distance of 80 cm in a quiet and dimly lit room.

4.1.3. Stimuli

Stimuli consisted of 3D rendered illustrations of twelve different indoor rooms. All the images were created with open source interior design software (Sweet Home 3D, version 4.6, eTeks Paris, France. Available at http://www.sweethome3d.com. See Appendix) and represented a bathroom, a bedroom, a cinema room, a children's room, a garage, a recreation room, a kitchen, a library, a living room, a music room, an office and a studio. The rooms were created balancing the colors of the environment and the number of objects in the room, in order to avoid any pop-out effect of the target. Each type of room was presented in a singular exemplar. The scenes were displayed at a resolution of 1200x900 pixels, subtending a visual angle of 22.83x17.12 degrees. Each scene included a yellow cup, which constituted the target in the visual search task. The cup's position was equally allocated to one of the six equal-sized rectangular segments in which the scene was divided, three (left, middle, right) above respectively below the horizontal midline. The cup was always presented in a meaningful position (e.g. cup on a surface and not floating in the air) and clearly visible.

4.1.4. Design and procedure

During the course of the first part of the experiment participants viewed 72 stimuli, divided into six blocks of twelve trials each. Each block contained six scenes in which the position of the target was repeated throughout the experiment (i.e. the repeated condition), in the remaining six scenes the target was presented in a different segment at each block (i.e. the non-repeated condition). Participants were not instructed that in some scenes the target’s position was repeated. Participants were asked to search for the yellow cup in each scene and to indicate the side of the handle by pressing the left or the right button of the mouse as quickly as possible. The side of the handle varied randomly, both in novel and repeated displays, so that no association of scene and response could be learnt. For each participant, the sequence of the stimuli was randomly assigned and the rooms that constituted the repeated displays were randomly drawn from the pool of all the rooms.

At the beginning of each trial, a fixation cross was presented at the center of the screen for 1000 ms in black on a gray background. Participants were instructed to fixate the cross. After the fixation cross, a blank gray screen was presented for 500 ms and after that, the search display was shown and remained visible until the response. Upon identifying the target, the subjects had to press the left button of the mouse when the cup had the handle on the left side or the right button of the mouse when the cup had the handle on the right side. After the response a feedback sound was provided (a high pitch tone for the correct response and a low pitch tone for the wrong response). An example of a trial sequence is displayed in Figure 14.

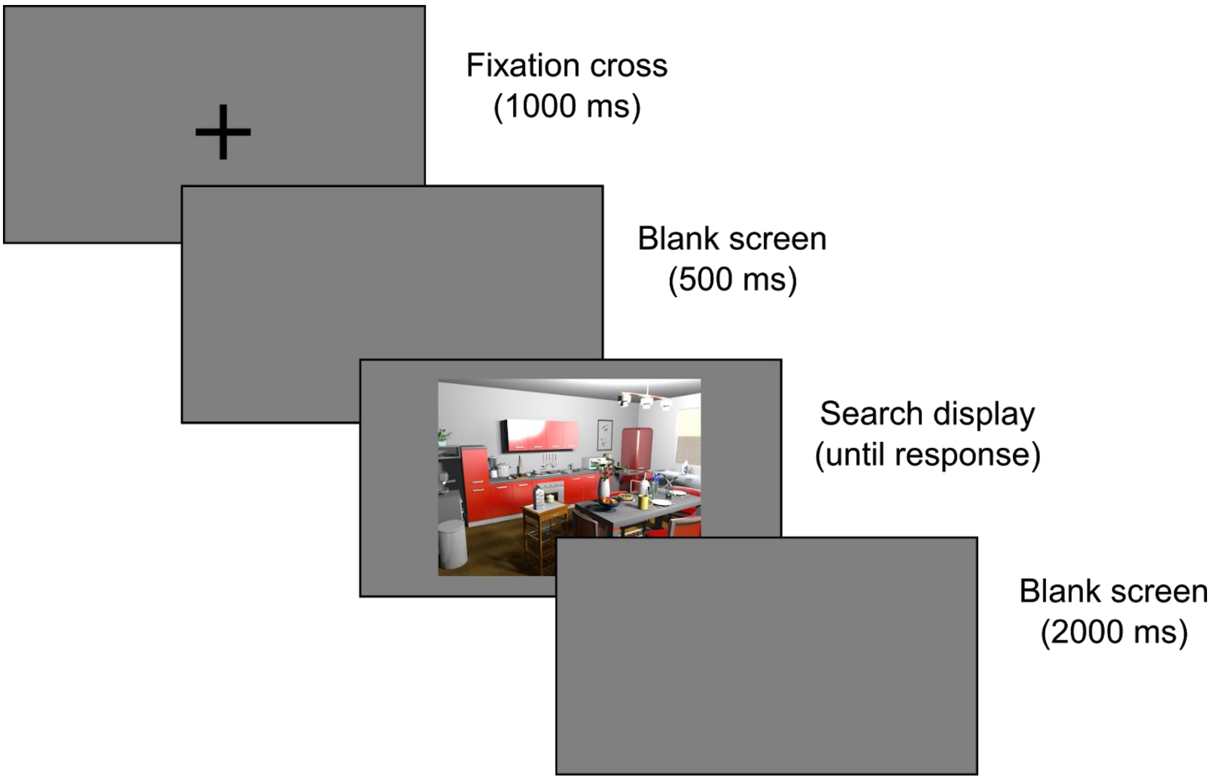


Figure 14. Example of a visual search trial.

In the second part of the experiment, participants completed a recognition/memory task. After a short break, the subjects performed a recognition task that contained the six real-world

3D rendered rooms that were presented with repeated target locations. The scenes were now rendered without the search target (i.e. the yellow cup) and participants were asked to move the mouse cursor and click on the previous target location. At the beginning of each trial, a fixation cross was presented at the center of the gray background for 1000 ms and then the scene was displayed until the response. No feedback was provided.

4.2. RESULTS

4.2.1. Visual search

Accuracy. In both groups the accuracy during the search task was high, on average, in the patient group the accuracy was 96.4%, whereas in the control group it was 98.5% ($t(20)=1.76$, $p=0.09$), indicating that the two groups did not present a significant difference in the response accuracies.

Response time. For statistical analysis log transformation was performed on the data and a three-way mixed-design ANOVA was run with group (AMD patients vs age-matched control group) as between-subjects factor, and configuration (repeated vs non-repeated) and epoch (1-3) as within-subjects factors (Figure 15). Overall, the main effect of Group was significant ($F(1,25)= 14.02$, $p< 0.001$), reflecting the patients' longer response time (control group mean= 1600 ms; patients mean= 2900 ms). In the course of the experiment, all the participants showed general learning skills (main effect of Epoch: $F(2,50)=11.50$, $p< 0.001$). As suggested by the main effect of Configuration, the scenes with constant target location showed a significant improvement ($F(1,25)= 9.24$, $p=0.005$). No other significant effect emerged. Furthermore, there was no reliable interaction between Group and other factors, suggesting the two groups produced a not significant different pattern in contextual cueing effect ($Fs>1.19$, $ps>0.312$).

However, the patients were generally older than the control group. In order to control for the potential age effects, all patients of 80 years or older (n=6) were removed from the experimental group. The mean age of the selected patient group (n=9) was 75.44 years, therefore it became comparable to the mean age of the controls (72.67 years; $t(18)=1.85$, $p=0.08$). After making the groups demographically comparable, a three-way mixed-design ANOVA was run on the age-selected group of patients and the controls. A significant effect of Group emerged ($F(1,25)= 18.54$, $p< 0.001$), reflecting the longer search time in the patients group compared to the healthy controls. That is, response time was faster in the control group (1600 ms) compared to the AMD patients group (2710 ms). Moreover, the significant main effect of Epoch ($F(2,50)=22.37$, $p< 0.001$) indicates that both groups showed a general improvement over time and the significant main effect of Configuration ($F(1,25)= 18.10$, $p< 0.001$) suggests that the search time was faster in the repeated displays. No interaction was significant ($Fs<2.73$, $ps<0.071$). Critically for the contextual cueing paradigm, the interaction Epoch x Configuration did not reach the significance ($F(1,25)=0.6$, $p= 0.552$). The lack of a significant Configuration x Epoch interaction led us to inquire if one condition was consistently easier than the other. Given the expectation that the two conditions presented the same difficulty at the beginning and only the repeated displays would have a considerable benefit at the end, an additional one-tailed t-test was carried out to compare the repeated and non-repeated condition in the first epoch and in the last epoch. There was no significant difference in the two conditions in the first epoch ($t(51)=1.21$, $p=0.117$), however, in the last epoch, the repeated condition showed a selective improvement ($t(51)=1.80$, $p=0.039$). The two groups did not present a different magnitude of contextual cueing effect (Group x Epoch x Condition: $F(2,50)=0.03$, $p=0.968$).

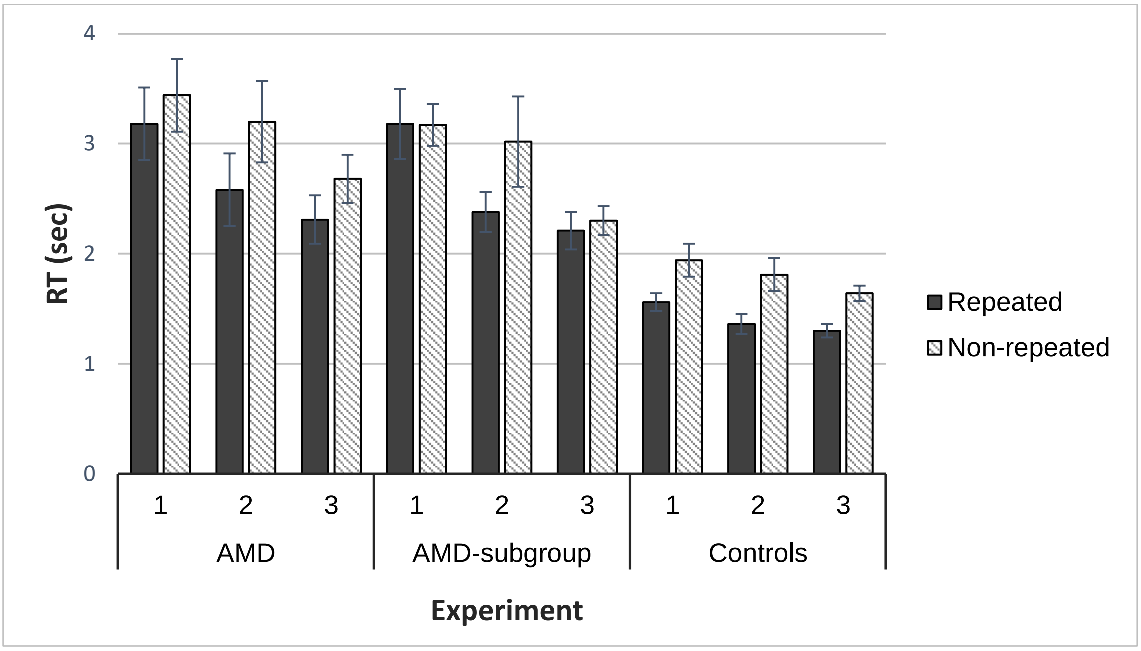


Figure 15. Mean of the reaction times (RTs) and standard errors for the repeated and non-repeated condition as a function of epoch (1-3), respectively in patients, in the group of AMD patients younger than 80 and in the control group.

4.2.2. Recognition test

For the purposes of the recognition test, displays were divided in six equal sized areas, each area contained the possible target locations. To be considered correct, the position (i.e. x,y coordinates given by pressing the mouse button) selected by the subject had to fall within the segment where the cup was presented in the main experiment. One subject in the AMD group was not included in the analysis because of a technical problem that led to data loss.

The average recognition accuracy in the full group of patients was 39.10%, whereas the accuracy in the demographically comparable group was 41.50%. To test whether the participants were able to recall the position of the target, we compared the accuracy obtained in the sample with the chance level value of 16%, since in each trial the probability to select the right position was 1/6. Patients' accuracy was significantly above chance in the full group ($t(13)=4.06$, $p=0.001$), as well as in the subgroup of patients ($t(7)=3.63$, $p=0.008$), indicating that they recalled the repeated target locations.

In the age-matched controls the accuracy was 48.25%. This, too, was significantly above chance ($t(11)=5.11$, $p<0.001$).

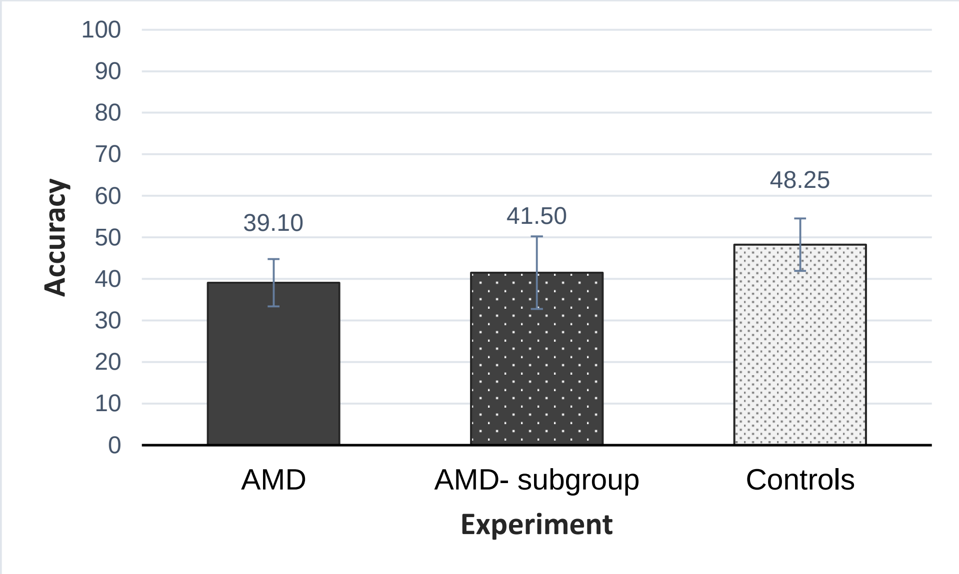


Figure 16. Accuracy in the recognition test respectively in the AMD group, in the AMD subgroup and in the age-matched controls.

Furthermore, neither the accuracy in the AMD group and in the control group ($t(23)=1.08$, $p=0.29$), nor the accuracy in the AMD sub-group and in the controls ($t(16)=0.71$, $p=0.485$) was significantly different (Figure 16).

4.3. DISCUSSION

The current work aimed to investigate whether the efficiency of visual search in repeated configurations is prevented in patients suffering from central vision loss compared to age-matched healthy controls. To do this, the performance of a group of patients suffering from AMD and a group of age-matched controls was tested in a visual search task where the stimuli were 3D-rendered real-world images. The main finding was that the search facilitation in the repeated displays was preserved in the patients as well as in the healthy controls. Overall, both groups showed an improved search performance, locating the target faster when it was kept in the same position across repetitions, in accordance with contextual cueing effect. It is reasonable to argue that the average benefit of 944 ms in the last epoch for the repeated configurations has not to be ascribed to a different difficulty of the conditions but instead to the facilitation given by the repeated exposure to the displays, as confirmed by a comparison between the search time in the repeated displays and in the novel displays at the beginning and end of the experiment. These results provided evidence that the foveal vision loss may have impaired the speed of the visual search, but it did not prevent contextual cueing effects. The results are consistent with previous studies that showed that contextual cueing was preserved from aging (Howard Jr, Howard, Dennis, Yankovich, & Vaidya, 2004; Geringswald et al., 2013, Merrill, Conners, Roskos, Klinger, & Klinger, 2013). In particular, Howard and colleagues' study (2004) demonstrated that a group of healthy elderly showed significant learning skills in finding the target faster in the repeated configuration than in the new ones. The performance was not significantly different than the one of a group of younger participants (i.e. college students), therefore the contextual cueing effect was preserved with ageing. Conversely, when the same task was proposed to patients with central visual impairment (i.e. AMD), they did not show any advantage by the repeated exposition to the displays, leading to the assumption that foveating the target was a basic requirement for developing contextual cueing (Geringswald et al., 2013). The critical factor can be ascribed to the nature of the stimuli, that is, all previous studies that investigated contextual cueing used the classical paradigm (Chung & Jiang, 1998) where the stimuli were arbitrary arrays of letters. For the first time, the idea that using a rich semantic environment can help the construction of contextual information despite the foveal impairment was tested. Thus, in contrast with previous research that tested contextual cueing in natural and simulated central vision loss (Geringswald et al., 2013; 2012), this study brought evidence that the foveation is not a critical requirement for learning the repeated association of scene and target-position, although performance for visual search was slower in AMD patients compared to the healthy controls. The major difference in this paradigm compared to the classic one resides in the richness of the semantic content in the 3D-rendered scenes, that quickly allows the categorization of the scene and permitted the participants to search more efficiently for the target (Brockmole & Henderson, 2006).

In real-world scenes, we can perceive a large amount of visual information without actually making a lot of eye movements because the initial scene representation is based on global properties of the image, more than on the detailed identification of the objects it contains (Brockmole, Castellano & Henderson, 2006), Therefore in natural scenes, it is possible to accomplish a visual search task with low-resolution peripheral vision. As suggested in van Asselen and Castelo-Branco (2009), the peripheral vision plays an important role in creating and storing the information regarding the target and its context. In agreement with previous findings, foveating a target was not necessary for processing the elements in the visual field, especially when, like in this experiment, the scenes provided a rich semantic context. A possible explanation lies in the fact that the recognition of some objects depends on the analysis of fine details, like for faces or words, whereas for other objects, such as buildings or scenes, the integration of visual information is based on peripheral areas (Levy, Hasson, Avidan, Hendler, & Malach, 2001). Therefore, contextual information in naturalistic scenes was not compromised in AMD patients.

Contextual cueing effect was attributed to a form of explicit learning because the participants were able to accurately indicate the correct location of the target, once it was removed from the scenes. The nature of the stimuli used in the experiment helped the creation of declarative knowledge that allowed the participants to memorized the repeated configurations (Standing, 1973), in line with the previous studies that studied contextual cueing in realistic scenes (Brockmole & Henderson, 2006; Brockmole et al., 2006). Moreover, the explicit learning proved to be preserved from aging. The scene-target location was explicitly encoded in memory, so that the participants from both groups performed above chance level when asked to indicate the correct location of the target in the repeated displays.

In conclusion, the findings suggest that patients with AMD, that is, suffering from foveal vision loss, exhibit high accuracy in visual search in realistic scenes. The search facilitation in the repeated scenes was preserved in the AMD patients as well as in the controls, despite the fact that the time spent to find the target was longer for the patients group compared to the healthy controls. Also, the facilitation for the repeated configurations resulted from an explicit learning of the spatial contingencies, thus participants could indicate the exact location of the target in the repeated scenes.

5

GENERAL DISCUSSION

I. SUMMARY OF FINDINGS

How visual perception operates over complex real-world scenes is an important issue for several cognitive sciences, psychology, neuroscience and machine learning. Contextual cueing is a useful paradigm for studying how contextual information is encoded and how it influences visual processing. I conducted three experiments that combined gaze-contingent technique with contextual cueing. Contextual cueing was here examined in realistic scenes that resembled interior rooms. The aim of the dissertation was to investigate the role of foveal/local and peripheral/global information in visual search with real-world scenes and the retrieval of the learnt configurations with respect to full vision search. To assess whether central or peripheral vision had the most impact on contextual cueing, a gaze-contingent display was simulated during the viewing of the scenes. In this way it was possible to dissociate the information processed by foveal vision from that acquired by the periphery. Therefore, the gaze-contingent could be either a window (Experiment 1) or a scotoma (Experiment 2). The window limited the visible area to 5° radius, making only the central information available during the search. The reverse condition was obtained by a gaze-contingent scotoma with the same size, thereby, the central vision was blocked by a gray patch and only the peripheral field was visible. Searching for a target in the presence of a simulated window or scotoma was an uncommon situation because it forced inexperienced participants to rely exclusively on peripheral or foveal information. This artificial way of exploring the scene required the participants to adapt their behaviour in order to control eye-movements. It can be assumed that the top-down controlled exploration of the display might have complicated the search guidance and ultimately affected the performance. For this reason, contextual cueing was then tested in a group of patients suffering from age-related macular degeneration (AMD), a clinical condition that leads to a progressive foveal vision loss (Experiment 3). For all the three studies it has been predicted that if contextual cueing in real-world scenes is based on local context, foveal vision should be sufficient for the effect to develop with a gaze-contingent window. On the contrary, if global context is necessary to guide visual search, exploring the scene with a simulated or natural scotoma should not impair contextual cueing effects. Alternatively, both manipulations may inhibit contextual cueing, suggesting the need of both local and global context to facilitate visual search.

By using a gaze-contingent window, at the end of a successful trial, participants only perceived the elements neighbouring the target, while the global context of the scene was darkened. Nevertheless, participants completed the search faster in the repeated configurations, that is, when the position of the target was constant across repetitions, compared to the configurations in which the target was in a new position at each presentation. Thus, contextual cueing fully developed over the course of the experiment when the search was performed with tunnel vision. The magnitude of the effect was not significantly different compared to that obtained accessing the entire visual field. The findings suggest that repeated exposure to the local information neighbouring the target facilitated visual search. Local context information was, therefore, sufficient for guiding attention to the relevant spot in the scene. The memory representation of the scene interacted with visual attention to enable efficient search in repeated configurations. Moreover, the findings underline the explicit nature of contextual cueing in real-world scenes. When later asked to recall the position of the target in repeated displays, participants proved to be able to retrieve the learnt context explicitly, by selecting the correct location of the target. The accuracy in the retrieval was not significantly different than when the scene was explored with full vision. The central scotoma allowed the investigation of the reverse situation. The gaze-contingent scotoma prevented the central part of the visual field to be processed, although participants were able to receive information from the periphery, thereby, accessing the global context of the scene. The results showed that when the configurations were repeated visual search became faster over the course of the experiment. The learning of the contingencies between scene and target position took place, thus indicating contextual cueing. Detailed foveal vision was not necessary for contextual learning to occur in real-world scenes, because global context enabled learnt configurations to guide visual search efficiently to the relevant portion of the display. Participants were shown to be sensitive to global context and to use this cue to locate the target. Contextual cueing was found to be explicitly encoded in memory, that is participants performed above chance level when asked to indicate the location of the target in the repeated displays. However, the loss of central vision interfered with the quality of the memory trace. The accuracy in the scotoma group was indeed significantly lower compared to the group of participants that explored the scene with full vision, indicating that participants struggled in the retrieval of the target position, compared to the control group. The gaze-contingent scotoma resembles a clinical condition called age-related macular degeneration (AMD), in which patients progressively lose central vision. Investigating contextual effect with natural central vision loss allowed for the examination of search behaviour in real-world scenes in an ecological manner. Patients with AMD exhibited high accuracy in finding the target in realistic scenes, suggesting that the central vision loss did not compromised visual search. Moreover, according to contextual cueing effect, the search facilitation in repeated scenes was preserved in the AMD patients as well as in the group of healthy age-matched controls, that is, the search time in finding the target in repeated configurations was faster than in non-repeated configurations. This evidence suggests that foveal vision loss did not compromise the ability of creating the association scene-target position, therefore foveating the target was not necessary for contextual cueing to occur. The facilitation for the repeated configurations resulted from an explicit learning of the spatial contingencies, that is, participants were able to recall the location of the target in the repeated scenes. Memory for the repeated contextual information efficiently guided visual search to the target position and the retrieval of the learnt context was explicitly encoded in both the AMD patients and the healthy controls. Additionally, the accuracy in the retrieval was not significantly different in the AMD patients compared to the healthy age-matched controls.

II. DISCUSSION OF FINDINGS IN RELATION TO PREVIOUS LITERATURE

Intelligent behaviour requires efficient selection of task relevant information. For this reason the human visual system became extremely efficient at extracting statistical regularities in the environment. In particular, the visual system adapted to be sensitive to repeated spatial locations in order to enhance performance on visual search task. This effect is known as contextual cueing (Chun & Jiang, 1998). The results presented here support previous research that investigated contextual cueing in real-world scenes. The repeated exposure to realistic environments led to a decreased search time for targets located consistently in the same position (Brockmole & Henderson, 2006). Observers encoded and retained the stable association scene-target location over repetitions and the memory for the repeated cue ultimately efficiently guided attention to the region that was relevant for the task. As a result, search time was faster for the repeated configurations compared to the new ones. In comparison with visual search in arrays of letters (Chun & Jiang, 1998), contextual cueing in real-world scenes occurred faster and was explicitly encoded in memory, so that the observers were aware of the repetitions and retrieved the position of the target with significantly high accuracy. For the learning of the contingencies to occur, the visual system matches the incoming display with previously acquired memories: the easier the match, the faster the search for relevant targets. However, it has been demonstrated (e.g. Brady & Chun, 2007; Olson & Chun, 2002; Song & Jiang, 2005) that a perfect match of the displays is not necessary for visual search to benefit from the repeated context; search time for partially matched displays is faster than that for new displays, although slower than when exact matching display is presented (Song & Jiang, 2005). Consistent with previous findings (Zang et al., 2010), contextual cueing was observed even when the scene was analyzed through a gaze-contingent display that excluded part of the visual scene.

In contextual cueing, target location can therefore be learnt relative to a fraction of the entire display (Brockmole et al., 2006 Experiment 2; Olson & Chun, 2002). Previous studies have debated which precise portion of the display guides visual search the most. Two contradictory views emerged regarding the nature of the information functional to guide search through repeated displays. Location of the task-relevant information was learnt relative to local context immediately neighbouring the target in letter arrays (Brady & Chun, 2007; Olson & Chun, 2002), while global context guided search in repeated natural scenes (Brockmole et al., 2006; Rosenbaum & Jiang, 2013). To solve the complicated issue of defining local and global dominance, a gaze-contingent technique was applied in the current work, with the purpose of dissociating the central/local context from the peripheral/global context.

The findings of the present work suggest that spatial contextual learning is encoded with respect of a hierarchical representation of the scene. That is, memory for the target location is learnt relative to optimal sub-regions of the scene. Visual search is optimally driven by the local context around the target, consistently with the evidence introduced by Brady and Chun (2007) and Olson and Chun (2002). Therefore, foveal vision contributes to elicit contextual cueing as much as having access to the entire visual field. However, when the local information is not available, learning of the repeated target position can take place in periphery, although the effect proved to be weaker. This consideration is consistent with the results of Olson and Chun (2002; Experiment 3), that showed that contextual cueing can develop across distance when target and context are separated by empty space, with no interfering information, similar to the case of a scotoma simulation. Therefore, the global recognition of the scene can take over in guiding visual search to the target, as proposed by Brockmole and colleagues (2006) and Rosenbaum and Jiang (2013). For example, when searching for a spoon, a person would preferably recall the particular drawer in which the spoon is usually found. If that information is not available, the global context "kitchen" is then used to guide visual search, based on previous knowledge. Access to the location of an object is thus subject to a hierarchical representation of the environment (Hirtle & Jonised, 1985; Taylor & Tversky, 1992), from the optimal to the sub-optimal regions of the scene.

Processing the local information around the target can be ecologically more efficient compared to encoding the entire display. In everyday life, paying attention to the local context neighbouring the target can be more opportune than attending the large scene, because objects may covary more frequently with items in their local context, compared to items in the long-range context (Brady & Chun, 2007). Furthermore, statistical learning, such as contextual cueing, is an intensive process to apply to the entire visual field with high efficiency. Encoding all the spatial relations among objects can be demanding, so in terms of efficient behaviour is more advantageous to learn the contingencies restricted to the local context that brought a successful outcome (i.e. finding the target). Nevertheless, when there is no information nearby the target, learning is extent to global context (Olson & Chun, 2002). In this case, peripheral vision, enabling the identity of the scene, guides visual search. A scene is a semantically coherent view of a real-world environment with a background and multiple elements arranged in a spatially licensed manner (Henderson & Hollingworth, 1999). Therefore, the scene identity provides intrinsic semantic constraints (Biederman et al., 1982) that prompt the range of plausible objects that a person can expect to find (Palmer, 1975). Moreover, global context suggests the position the objects occupy in a scene and the relation they have with respect to each other (Biederman et al., 1982; Hollingworth, 2006). The importance of global context is especially relevant in naturalistic scenes and it does not hold in artificial displays, where the items are placed in an homogeneous, meaningless space. Thus, global context supplies a rich amount of information about the scene and provides knowledge about the structure of the scene and the relation among its parts. Global context allows observers to activate the schema correspondent to the scene, which facilitates the analysis of the objects arranged in the scene in a consistent manner (Hollingworth, 1999). Furthermore, global context provides the recovery of knowledge from past experiences faster than a small subset of elements placed around the target, allowing a quicker transfer of learning (Brockmole et al., 2006). This top-down mechanism for guiding attention does not need to use detailed foveal perception, but instead relies on peripheral vision (Zang et al., 2015). Indeed, some objects, such as buildings and scenes, are better analyzed over larger retinal distances (Levy et al., 2001). Using peripheral parts of the visual field for exploring visual scenes becomes crucial in situation of natural foveal vision loss. When central vision is unavailable, visual processes need to be carried out by periphery. Patients with age-related macular degeneration (AMD) usually develop an eccentric extrafoveal retinal spot outside the scotoma, called preferred retinal locus

(PRL), that acts like a pseudo-fovea (Maniglia, Soler, Cottereau, & Trotter, 2018; Timberlake, Peli, Essock, & Augliere, 1987; von Noorden & Mackensen, 1962). Therefore, patients with AMD learn to explore the environment with the extrafoveal part of the retina, that was not damaged by the macular degeneration, in order to compensate for the lack of central vision. The PRL allows patients to relocate visual attention to the peripheral part of the visual field for encoding the objects and their location. The target position is then stored in visual long-term memory. Explicit visual long-term memory in patients suffering from central vision loss has been found to be as accurate as in the healthy age-matched controls (Geringswald, Herbik, Hofmüller, Hoffmann, & Pollmann, 2015). These results are confirmed by the recognition task in Experiment 3. Therefore, the findings suggest that the PRL contributes to the deployment of attention to the periphery, in order to encode the objects that are not foveated due to the scotoma. The encoded context is then stored in visual long-term memory and it is used to guide attention to the relevant portion of the display. Over repetitions the repeated exposure to the learnt context enhances visual search. Patients with central vision impairment require several years to develop the PRL and use peripheral vision efficiently. Covertly shifting attention to encode objects with low resolution in the peripheral part of the visual field, has several disadvantages. Therefore, when inexperienced normal-sighted observers are challenged with a gaze-contingent scotoma, visual long-term memory for objects is compromised (Geringswald, Porracin, & Pollmann, 2016), as revealed by the results of the recognition task in Experiment 2. The compromised accuracy in retrieving the location of the target may be the reason for the relatively weak contextual cuing effect obtained with the scotoma simulation. On the contrary, foveating the objects allows the retrieval of a relatively detailed representation of the scene (Hollingworth & Henderson, 2002). Considering that foveation of the objects is closely related to the allocation of visual attention, the attended information is therefore stored in visual long-term memory (Deubel & Schneider, 1996). This is in line with the results of the recognition task performed after searching with tunnel vision (Experiment 1), in which the observers recalled the location of the target with accuracy not significantly different than the observers that inspected the scene using the entire visual field.

III. CONCLUSIONS, LIMITATIONS AND FUTURE DIRECTIONS

In conclusion, the findings presented here converge with the idea that currently encountered information is linked to information repeatedly encountered in the past (Song & Jiang, 2005). The human visual system can retrieve previous memories on the basis of a partially matching display (Brady & Chun, 2007; Olson & Chun, 2002), however the quality of the memory trace is conditioned by the capability of processing the central part of the visual field. Peripheral vision reaches the highest efficiency in scanning the environment when the person develops a PRL, that acts as a pseudo-fovea. However, contextual cueing can also be developed in conditions of simulated vision loss, both peripheral and central. The present dissertation aimed to attempt to reconcile the contradictory views regarding the nature of the information that guides visual search in repeated scenes. For this purpose, gaze-contingent techniques were applied, so that the role of peripheral/global and central/local context could be selectively investigated. Compared to previous studies in real-world scenes, in which local and global information were continuously available during the search, the present investigation found a way to fully separate the role of foveal and peripheral context. The use of gaze-contingent displays introduces a novel element in the studying of contextual cueing in real-world scenes, and serves the purpose of examine the ability to extract and remember consistent spatial relations in real-world environments, when the search is affected by either foveal or peripheral vision loss.

Due to the clinical population involved in Experiment 3, the results are based on a small number of participants. Moreover, the progress of the AMD is not homogeneous among patients. A bigger sample might in the future extend the understanding of the effects of macular degeneration in visual search. For example, studying the search performance in different stages of macular degeneration could clarify whether there is actually a linear correlation between the ability of using efficiently peripheral vision and the strength of contextual cueing effect. Future studies might additionally investigate contextual cueing effects in real-world scenes in another condition of natural visual loss. Retinitis pigmentosa, for example, is a disorder that affects the retina, it consists of the progressive loss of sight in the peripheral part of the visual field, so that patients experience a natural tunnel vision.

In summary, the dissertation offers evidence that contextual cueing in real-world scenes can develop when the scene is scanned through a gaze-contingent display that excluded part of the visual scene. In particular, foveal vision contributes to elicit contextual cueing as much as having access to the entire visual field, the local elements neighboring the target guide efficiently visual search in repeated displays. However, when the local information is not available, learning of the repeated target position can take place in periphery. Thanks to the global recognition of the scene, the scene’s identity is enabled and therefore the contingencies between the elements. Even though in this latter case, the effect proved to be weaker. Spatial contextual learning is therefore encoded with respect to a hierarchical representation of the scene. That is, memory for the target location is learnt relative to optimal sub-regions of the scene. Foveating the objects allows for the retrieval of a relatively detailed representation of the scene, as demonstrated by the use of the gaze-contingent window. Instead, visual long-term memory for objects is compromised when inexperienced normal-sighted observers are challenged with a gaze-contingent scotoma. However, when patients with age-related macular degeneration, that is with a natural scotoma, develop a PRL outside the damaged portion of the retina, they can encode the objects and their location into visual long-term memory, using the peripheral part of the visual field. The deployment of attention to the periphery contributes to efficiently encode the target location in visual search tasks and therefore it speeds up the search.

ACKNOWLEDGEMENTS

There have been several people who have walked alongside me during these years.

I would first like to express my gratitude to my advisor Prof. Dr. Pollmann, for the support throughout the course of my work.

My sincere thanks go to my fellow labmates, for always being willing to help me in this journey.

In addition, I am grateful to Stefanie Senf, who acquired the data presented in Experiment 3.

Finally, I would like to thank all the people who did believe in me, for giving me the motivation to arrive at the last world of this thesis.

Eleonora

REFERENCES

Ambati, J., & Fowler, B. J. (2012). Mechanisms of age-related macular degeneration. *Neuron*, *75*(1), 26-39.

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.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485-496.

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*(7), 460-466.

Beard, B. L., & Ahumada, A. J. (1999). Detection in fixed and random noise in foveal and parafoveal vision explained by template learning. *JOSA A*, *16*(3), 755-763.

Biederman, I. (1972). Perceiving real-world scenes. *Science*, *177*(4043), 77-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143-177.

Boucart, M., Despretz, P., Hladiuk, K., & Desmettre, T. (2008). Does context or color improve object recognition in patients with low vision? *Visual Neuroscience*, *25*(5-6), 685-691.

Boucart, M., Dinon, J. F., Despretz, P., Desmettre, T., Hladiuk, K., & Oliva, A. (2008). Recognition of facial emotion in low vision: a flexible usage of facial features. *Visual Neuroscience, 25*(4), 603-609.

Boucart, M., & Tran, T. H. C. (2011). Object and Scene Recognition Impairments in Patients with Macular Degeneration. In *Object Recognition*. InTech.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226(5241)*, 177-178.

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.

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.

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

Brockmole, J. R., & 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.

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), 1406-1418.

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *The Journal of Physiology*, *181*(3), 576-593.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170-178.

Chun, M. M. (2003). Scene perception and memory. *Psychology of Learning and Motivation*, *42*, 79-108.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73-101.

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

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

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

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50(1)*, 243-271.

Chun, M. M., & Turk-Browne, N. B. (2008). Associative learning mechanisms in vision. In S. J. Luck & A. R. Hollingworth (Eds.), *Visual Memory*, (pp 209-245). Oxford: Oxford University Press.

Chun, M. M., & Wolfe, J. M. (2001). Visual attention. In B. Goldstein (Ed.) *Blackwell handbook of perception*, pp 272-310.

Coletta, N. J. Williams, D. R. (1987). Psychophysical estimate of extrafoveal cone spacing. *Journal of the Optical Society of America A, Optics and Image Science*, 4, 1503–1513.

Cornelissen, F. W., Bruin, K. J., & Kooijman, A. C. (2005). The influence of artificial scotomas on eye movements during visual search. *Optometry & Vision Science*, *82*(1), 27-35.

De Jong, P. T. (2006). Age-related macular degeneration. *New England Journal of Medicine*, *355*(14), 1474-1485.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827-1837.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Attention, Perception, & Psychophysics*, *14*(1), 155-160.

Fine, I., MacLeod, D. I., & Boynton, G. M. (2003). Surface segmentation based on the luminance and color statistics of natural scenes. *JOSA A*, *20*(7), 1283-1291.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *PNAS*, *99*(24), 15822-15826.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316-355.

Geringswald, F., Baumgartner, F., & Pollmann, S. (2012). Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Frontiers in Human Neuroscience*, *6*(134), 1-9.

Geringswald, F., Baumgartner, F. J., & Pollmann, S. (2013). A behavioral task for the validation of a gaze-contingent simulated scotoma. *Behavior Research Methods*, *45*(4), 1313-1321.

Geringswald, F., Herbik, A., Hoffmann, M. B., & Pollmann, S. (2013). Contextual cueing impairment in patients with age-related macular degeneration. *Journal of Vision*, *13*(3), 1-18.

Geringswald, F., Herbik, A., Hofmüller, W., Hoffmann, M. B., & Pollmann, S. (2015). Visual memory for objects following foveal vision loss. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1471-1484.

Geringswald, F., & Pollmann, S. (2015). Central and peripheral vision loss differentially affects contextual cueing in visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1485-1495.

Geringswald, F., Porracin, E., & Pollmann, S. (2016). Impairment of visual memory for objects in natural scenes by simulated central scotomata. *Journal of Vision*, *16*(2), 1-12.

Green, D. G. (1970). Regional variations in the visual acuity for interference fringes on the retina. *The Journal of Physiology*, *207*(2), 351-356.

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524-533.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498-504.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*(1), 243-271.

Henderson, J. M., Weeks Jr, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210-228.

Hirsch, J., & Miller, W. H. (1987). Does cone positional disorder limit resolution?. *JOSA A*, *4*(8), 1481-1492.

Hirsch, J., & Curcio, C. A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research*, *29*(9), 1095-1101.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, *13*(3), 208-217.

Hoffman, J. E., & Nelson, B. (1981). Spatial selectivity in visual search. *Perception & Psychophysics*, *30*(3), 283-290.

Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, *14*(4-8), 781-807.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113-136.

Hollingworth, A. Schrock, G. Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, 29, 296–304.

Holmes, D. L. Cohen, K. M. Haith, M. M. Morrison, F. J. (1977). Peripheral visual processing. *Perception & Psychophysics*, 22, 571–577.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Attention, Perception, & Psychophysics*, *57*(6), 787-795.

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

Howard Jr, J. H., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, *18(1)*, 124-134.

Jager, R. D., Mieler, W. F., & Miller, J. W. (2008). Age-related macular degeneration. *New England Journal of Medicine*, *358*(24), 2606-2617.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review*, *12*(1), 100-106.

Jiang, Y., Song, J. H., & Rigas, A. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review*, *12*(3), 524-529.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—configuration or individual locations?. *Attention, Perception, & Psychophysics*, *66*(3), 454-463.

Klaver, C. C., Wolfs, R. C., Vingerling, J. R., Hofman, A., & de Jong, P. T. (1998). Age-specific prevalence and causes of blindness and visual impairment in an older population: the Rotterdam Study. *Archives of Ophthalmology*, *116(5)*, 653-658.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cuing by global features. *Perception & Psychophysics*, *68*(7), 1204-1216.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 1-16.

Legge, G. E., Ross, J. A., Isenberg, L. M., & Lamay, J. M. (1992). Psychophysics of reading. Clinical predictors of low-vision reading speed. *Investigative Ophthalmology & Visual Science*, *33*(3), 677-687.

Legge, G. E., Rubin, G. S., Pelli, D. G., & Schleske, M. M. (1985). Psychophysics of reading—II. Low vision. *Vision Research*, *25*(2), 253-265.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4(5)*, 533-539.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565-572.

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.

Maniglia, M., Soler, V., Cottereau, B., & Trotter, Y. (2018). Spontaneous and training-induced cortical plasticity in MD patients: Hints from lateral masking. *Scientific Reports*, *8*(90), 1-11.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Attention, Perception, & Psychophysics*, *17*(6), 578-586.

Merrill, E. C., Conners, F. A., Roskos, B., Klinger, M. R., & Klinger, L. G. (2013). Contextual cueing effects across the lifespan. *The Journal of Genetic Psychology, 174(4),* 387-402.

Midena, E., Degli Angeli, C., Blarzino, M. C., Valenti, M., & Segato, T. (1997). Macular function impairment in eyes with early age-related macular degeneration. *Investigative Ophthalmology & Visual Science*, *38*(2), 469-477.

Miller, W. H. (1979). Ocular optical filtering. In *Comparative Physiology and Evolution of Vision in Invertebrates* (pp. 69-143). Springer Berlin Heidelberg.

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1299–1313.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, *9*(3), 273-302.

Osterberg, G. A. (1935). Topography of the layer of the rods and cones in the human retima. *Acta Ophthalmol*, *13*(6), 1-102.

Palmer, T. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519-526.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*(1), 629-651.

Polyak, S. L. (1941). *The retina*. Chicago: University of Chicago Press.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 509-522.

Querques, G., Avellis, F. O., Querques, L., Bandello, F., & Souied, E. H. (2011). Age-related macular degeneration. *Clinical Ophthalmology, 5*, 593-601.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*(4417), 468-469.

Rayner, K. Inhoff, A. W. Morrison, R. E. Slowiaczek, M. L. Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 167–179.

Reder, S. M. (1973). On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behavior Research Methods*, *5*(2), 218-228.

Rodieck, R. W. (1998). *The first steps in seeing*. Sunderland, MA: Sinauer Associates, Inc.

Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception, & Psychophysics*, *75*(5), 888-899.

Schacter, D. L. (1994). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 233–268). Cambridge, MA: MIT Press.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156-163.

Shimozaki, S. S. Chen, K. Y. Abbey, C. K. Eckstein, M. P. (2007). The temporal dynamics of selective attention of the visual periphery as measured by classification images. *Journal of Vision*, 7, (12):10, 1–20.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261-267.

Sjostrand, J., & Friseu, L. (1977). Contrast sensitivity in macular disease. *Acta Ophthalmologica*, *55*(3), 507-514.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition, 36(2)*, 403-415.

Snyder, A. W., & Miller, W. H. (1977). Photoreceptor diameter and spacing for highest resolving power. *JOSA*, *67*(5), 696-698.

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)*, 322-330.

Standing, L. (1973). Learning 10000 pictures. *The Quarterly journal of experimental psychology*, *25*(2), 207-222.

Steinberg, R. H., Reid, M., & Lacy, P. L. (1973). The distribution of rods and cones in the retina of the cat (Felis domesticus). *Journal of Comparative Neurology, 148(2)*, 229-248.

Sterling, P., Freed, M. A., & Smith, R. G. (1988). Architecture of rod and cone circuits to the on-beta ganglion cell. *Journal of Neuroscience*, *8*(2), 623-642.

Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition*, *20*(5), 483-496.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Attention, Perception, & Psychophysics*, *51*(6), 599-606.

Timberlake, G. T., Peli, E., Essock, E. A., & Augliere, R. A. (1987). Reading with a macular scotoma. II. Retinal locus for scanning text. *Investigative Ophthalmology & Visual Science*, *28*(8), 1268-1274.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (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.

Tran, T. H. C., Rambaud, C., Despretz, P., & Boucart, M. (2010). Scene perception in age-related macular degeneration. *Investigative Ophthalmology & Visual Science*, *51(12)*, 6868-6874.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97-136.

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2015). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23(1)*, 87–102.

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

Velez-Montoya, R., Oliver, S. C., Olson, J. L., Fine, S. L., Quiroz-Mercado, H., & Mandava, N. (2014). Current knowledge and trends in age-related macular degeneration: genetics, epidemiology, and prevention. *Retina*, *34*(3), 423-441.

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1358-1371.

von Noorden, G. K., & Mackensen, G. (1962). Phenomenology of eccentric fixation. *American Journal of Ophthalmology*, *53*(4), 642-661.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Attention, Perception, & Psychophysics*, *56*(5), 495-500.

Wang, Y. Z., Wilson, E., Locke, K. G., & Edwards, A. O. (2002). Shape discrimination in age-related macular degeneration. *Investigative Ophthalmology & Visual Science*, *43*(6), 2055-2062.

Westheimer, G. (1960). Modulation thresholds for sinusoidal light distributions on the retina. *The Journal of Physiology*, *152*(1), 67-74.

Williams, C. C., Henderson, J. M. & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception & Psychophysics*, *67*(5), 816-827.

Williams, D. R. (1985). Aliasing in human foveal vision. *Vision Research*, *25*(2), 195-205.

Williams, D. R. (1986). Seeing through the photoreceptor mosaic. *Trends in Neurosciences*, *9*, 193-198.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202-238.

Wolfe, J. M. (1998). Visual search. *Attention*, *1*, 13-73.

Yamada, E. (1969). Some structural features of the fovea centralis in the human retina. *Archives of Ophthalmology*, 82, 151–159.

Yantis, S. (1998). Control of visual attention. *Attention*, *1*(1), 223-256.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(1), 95-107.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision* (pp. 171-211). Springer US.

Zang, X., Jia, L., Müller, H. J., & Shi, Z. (2015). Invariant spatial context is learned but not retrieved in gaze-contingent tunnel-view search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 807-819.

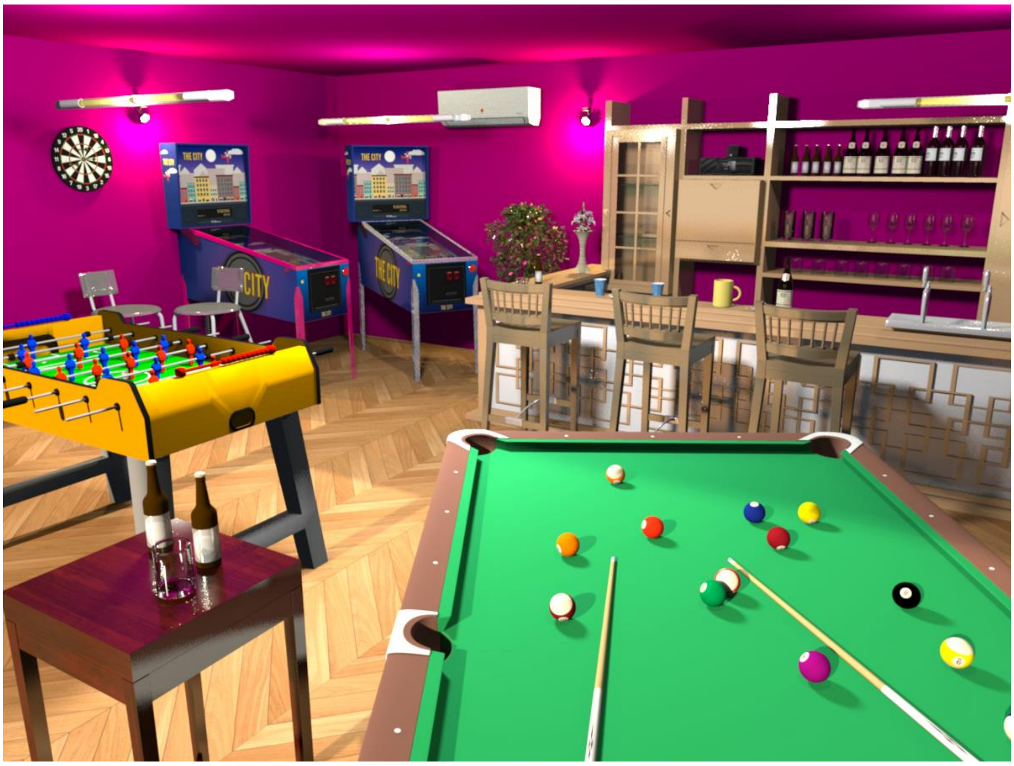
Zellin, M., von Mühlenen, A., Müller, H. J., & Conci, M. (2013). Statistical learning in the past modulates contextual cueing in the future. *Journal of Vision*, *13*(3), 1-19.

APPENDIX: SCENES

Bathroom

Bedroom

Cinema room



Game room

Garage

Children room

Kitchen

Library

Living room

Music room

Office

Studio

Division of the scene into six equal sized sections. Example of a scene used during the recognition task. The yellow lines were not visible to the participants at any time.

M.Sc., Eleonora Porracin
Annastraße 1
39108 Magdeburg

Erklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation zum dem Thema

“Role of foveal and peripheral vision in contextual cueing and its retrieval in real-world scenes”

selbständig verfasst, nich schon als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständing angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, 25.04.2018

M.Sc., Eleonora Porracin