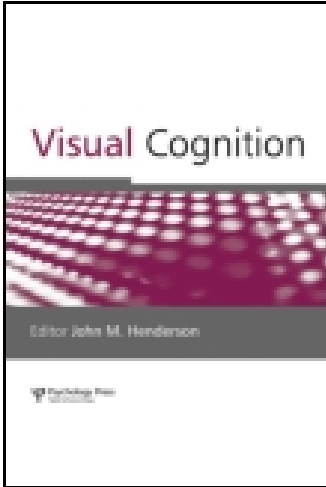


This article was downloaded by: [Eastern Michigan University]

On: 09 October 2014, At: 18:44

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



Visual Cognition

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pvis20>

A glimpse is not a glimpse: Differential processing of flashed scene previews leads to differential target search benefits

Melissa L.-H. Vö^{a b} & Werner X. Schneider^{c d}

^a Department of Psychology , Ludwig-Maximilians-Universität München , Germany

^b Psychology Department , University of Edinburgh , UK

^c Department of Psychology , Ludwig-Maximilians-Universität München

^d Department of Psychology , Bielefeld University , Germany

Published online: 11 Dec 2008.

To cite this article: Melissa L.-H. Vö & Werner X. Schneider (2010) A glimpse is not a glimpse: Differential processing of flashed scene previews leads to differential target search benefits, *Visual Cognition*, 18:2, 171-200, DOI: [10.1080/13506280802547901](https://doi.org/10.1080/13506280802547901)

To link to this article: <http://dx.doi.org/10.1080/13506280802547901>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should

not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

A glimpse is not a glimpse: Differential processing of flashed scene previews leads to differential target search benefits

Melissa L.-H. Võ

Department of Psychology, Ludwig-Maximilians-Universität München, Germany, and Psychology Department, University of Edinburgh, UK

Werner X. Schneider

Department of Psychology, Ludwig-Maximilians-Universität München, and Department of Psychology, Bielefeld University, Germany

What information can we extract from an initial glimpse of a scene and how do people differ in the way they process visual information? In Experiment 1, participants searched 3-D-rendered images of naturalistic scenes for embedded target objects through a gaze-contingent window. A briefly flashed scene preview (identical, background, objects, or control) preceded each search scene. We found that search performance varied as a function of the participants' reported ability to distinguish between previews. Experiment 2 further investigated the source of individual differences using a whole-report task. Data were analysed following the "Theory of Visual Attention" approach, which allows the assessment of visual processing efficiency parameters. Results from both experiments indicate that during the first glimpse of a scene global processing of visual information predominates and that individual differences in initial scene processing and subsequent eye movement behaviour are based on individual differences in visual perceptual processing speed.

Keywords: Scene perception; Attention; Eye movement control; Individual differences; Processing efficiency; Theory of Visual Attention.

Please address all correspondence to Melissa Le-Hoa Võ, Visual Cognition Unit, Psychology Department, 7 George Square, University of Edinburgh, Edinburgh, EH8 9JZ, UK. E-mail: melissa.vo@ed.ac.uk

We want to thank Carrick Williams as well as two anonymous reviewers for their helpful comments on a prior version of this manuscript. Further, we want to thank Ellen Matthias for her support and helpful comments regarding the application of the Theory of Visual Attention to the data of Experiment 2. This project was supported by a grant of the DFG excellence initiative research cluster "Cognition for Technical Systems-CoTeSys" DFG (De 336/2), and by the DFG-FOR 480 (TP 4).

Imagine visiting friends and helping out in their kitchen. When being asked to fetch some plates you probably are able to find them at a location that seems plausible to you without having to search the entire room. This seems trivial, but involves a number of cognitive processes. For example, you have to recognize the scene you are about to act in as being part of a kitchen. In order to fulfil the given task you have to further activate your implicit knowledge about kitchens, i.e., their typical layout, their functionality, and typical locations of typical appliances. The combination of these long-term representations with the currently evolving representation of the specific kitchen as well as with task knowledge, will in most cases lead to search benefits due to the active exploration of only those parts of the kitchen that have a high probability of containing the plates (see Torralba, Oliva, Castelano, & Henderson, 2006).

The study presented here investigated what information extracted from a first glimpse of a complex naturalistic scene can modulate the deployment of attention and eye movements during subsequent target search. We will see that the initial scene representation has significant influence on where we look next. However, a hardly discussed issue is that of individual differences in perceiving briefly presented complex scenes. People might differ in the way they benefit from an initial glimpse of a scene. Thus, we were also interested in the cognitive processes that differ between individuals during initial scene processing and whether these individual differences could then also affect attention allocation and eye movement control during subsequent search.

ATTENTION ALLOCATION AND EYE MOVEMENT CONTROL IN SCENE PERCEPTION

There is no doubt that when viewing a natural scene, attention and the human eye do not move around randomly. However, there has been a dispute regarding the degree to which eye movements during scene perception are influenced by bottom-up image properties such as contrast or colour on the one hand or by top-down factors such as the current task or scene knowledge on the other (for a review see Henderson, 2007). The first neurocomputational models of visual attention that dealt with natural scenes strongly relied on attention control by bottom-up image saliency (e.g., Itti & Koch, 2000; Itti, Koch, & Niebur, 1998; Parkhurst, Law, & Niebur, 2002). On the basis of combined information from different feature maps (e.g., colour, intensity, and orientation), highly salient regions of an image can be located which are assumed likely to attract observers' attention. These models perform quite well when no specific task is driving the observer's exploration of an image (Underwood, Foulsham, van Loon,

Humphreys, & Bloyce, 2006). However, in most real-world settings an observer's activity is influenced by a given task, as in our example, to find a target amongst a number of distractors. Underwood and colleagues (Underwood & Foulsham, 2006; Underwood et al., 2006) have tried to disentangle the specific contributions of bottom-up visual saliency and top-down task demands. In their experiments, participants inspected pictures of natural scenes in which two objects of interest were placed, one of which was characterized by high and the other by low visual saliency according to the Itti and Koch (2000) algorithm. The task was modified to determine whether visual saliency is invariably the dominant attractor of fixations, or whether task influences can provide a cognitive override that renders saliency secondary. When the participants were told to inspect the scene in preparation for a subsequent memory task, visually salient objects attracted early fixations in support of a saliency map model of scene inspection. However, when participants had to search for a specific target amongst distractors of higher visual saliency, the target would attract attention to a greater degree despite being less salient. These results support a version of the saliency map hypothesis in which task demands can cognitively override a purely bottom-up driven saliency map.

Recent computational models have taken the modulation of attention allocation by cognitive processes into account assuming the combined influence of both bottom-up and top-down information (see Navalpakkam & Itti, 2005; Torralba et al., 2006). In those models, bottom-up processing based on low-level image features interacts with the top-down processing of scene gist and spatial layout allowing a shift of attention and correspondingly the eyes to locations that have a high probability of containing the search target. In their contextual guidance model, Torralba and colleagues (2006) propose that an image is analysed in two parallel pathways: The local and the global pathway. Both pathways share the first stage during which the image is filtered by a set of multiscale-oriented filters. The local representation comprises each spatial location independently and is used to both compute local saliency peaks and perform object recognition. The global pathway, on the other hand, represents the entire image holistically by extracting global statistics from the image that allows the activation of knowledge and expectations regarding a specific scene—so-called scene priors. Task demands are supposed to influence only the global pathway in that it provides information on the expected location of the target as a function of scene priors. Thus, a key feature of the model is the interaction of local and global processing within the first glimpse in order to rapidly narrow down the search area to those parts of the scene that most probably contain the target.

THE INFLUENCE OF INITIAL SCENE REPRESENTATIONS

In order to exert its influence on subsequent eye movements, the initial scene representation has to be stored in visual memory across several saccades. During search, more detailed information is then continuously acquired with each fixation adding to the evolving scene representation (e.g., Henderson & Castelhana, 2005; Hollingworth, 2005; Tatler, Gilchrist, & Land, 2005; Tatler, Gilchrist, & Rusted, 2003). To investigate whether the initial scene representation acquired from a flashed preview of a scene can be stored in such a way that it continuously exerts its influence on visual search in a real-world scene, Castelhana and Henderson (2007) used the “Flash-Preview Moving-Window Paradigm”. This paradigm elegantly combines the brief tachistoscopic viewing method typically used in scene identification (or scene categorization) experiments with the moving window technique typically used to investigate eye movements under restricted viewing conditions.

In their study, participants were asked to search for target objects in scenes while their eye movements were recorded. Prior to presentation of the search scene, a scene preview was briefly presented for 250 ms. Then followed a word indicating the identity of the target, after which the search scene was presented. However, the search scene was only visible through a gaze-contingent moving window with a 2° diameter centred at fixation within a scene. This paradigm allows selective manipulation of the information provided by the preview of the scene, at the same time enabling the investigation of subsequent eye movement behaviour and controlling for the information uptake during the actual search by restricting the latter to foveal vision only. With the restricted view through the moving window during search, information uptake cannot be influenced by parafoveal vision, which is needed for the viewer to extract a vector of global features and rapidly set up scene priors. Possible preview benefits due to global processing must therefore be attributed solely to the memory-based scene representation formed as a result of the processing of the briefly flashed preview.

In a number of experiments, Castelhana and Henderson (2007) were able to show not only that the initial scene representation can be used to predict highly probable target locations, but also that this initial representation continues to be available in an abstract manner. For example, an identical scene preview led to significant search benefits during subsequent target search as compared to a different or meaningless scene preview. Also, a scene preview still benefited subsequent target search when it was identical, but minimized in its size as compared to the search scene. However, a preview did not benefit search when it sustained the conceptual category of the following search scene, while differing in its visual details. Thus, the initially

crude scene representation seems to be stored in an abstract manner, but needs more specific information (e.g., the particular spatial layout of the scene) in order to benefit target search.

INDIVIDUAL DIFFERENCES IN EYE MOVEMENT CONTROL

In the study presented here we used the “Flash-Preview Moving-Window Paradigm” (Castelhano & Henderson, 2007) to investigate what information presented during the first glimpse of a scene shows greatest benefits when subsequently having to search for target objects in naturalistic scenes. Additionally, we were interested in whether participants would differ in the ability to extract useful information from briefly flashed scene previews. Although Castelhano and Henderson (2007) showed that scene representations generated from only briefly presented scenes can generally influence later target search, the role of individual differences in initial scene processing and its influence on subsequent eye movements in visual search has not yet been addressed in greater detail.

There is some evidence that individuals differ in the way they process visual input as a function of expertise—qualitatively different ways of processing information as a result of experience—or a more general processing efficiency unrelated to the specific visual input. Interestingly, individual as well as cultural differences have been reported to also show in eye movement behaviour when viewing natural scenes (e.g., Andrews & Coppola, 1999; Castelhano & Henderson, 2008; Chua, Boland, & Nisbett, 2005; Nisbett & Masuda, 2003; Rayner, Li, Williams, Cave, & Well, 2007; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). For example, Underwood and colleagues (2003) have shown that scan paths during driving differ as a function of expertise, i.e., experienced drivers monitored other road users more often than novice drivers, who showed little ability to switch the focus of their attention as potential hazards appeared. Further evidence of individual differences in processing visual input comes from a recent study by Brockmole, Hambrick, Windisch, and Henderson (2008), who found that expert chess players developed a contextual cueing effect during target search, which was four times as large as the one generated by novices. In a change detection experiment using alternating displays with a presentation rate of 500 ms, Werner and Thies (2000) showed that domain-specific expertise increased the ability to detect changes for flashed scenes, implying that there are individual differences regarding rapid picture processing as a function of expertise as well. Apart from domain-specific expertise, studies on reading have shown that eye movement patterns differ between good and poor readers such that poor readers tend to fixate longer and make shorter saccades than good readers

due to differences in general processing efficiency (e.g., Eden, Stein, Wood, & Wood, 1994; Hutzler & Wimmer, 2004). Thus, there seems to be evidence that individuals greatly differ in their ability to process visual information. The present study was mainly concerned with the effects of the degree of initial scene processing rather than with the investigation of effects due to specific expertise. When the presentation time of a complex visual input is limited to a split second, the individual processing efficiency in particular might lead to differential benefits in using the flashed information for effective eye movement control.

An integrated theoretical and methodological approach, which permits the assessment of components of processing efficiency, is the Theory of Visual Attention (TVA; Bundesen, 1990, 1998). TVA assumes that a number of latent processes underlie overt performance. These processes are formally described by a coherent, mathematical theory in terms of a set of (mathematically) independent quantitative parameters (for a detailed mathematical description, see Bundesen, 1990, 1998; Kyllingsbaek, 2006). For example, the general efficiency of the visual processing system is reflected in the parameters visual perceptual processing speed C (number of visual elements processed per second) and visual short-term memory (VSTM) storage capacity K (number of elements maintained in parallel). Both parameters can be assessed using a whole-report task, in which participants are briefly presented with arrays of simple stimuli, e.g., letters at varying exposure durations from which they have to identify (name) as many as possible. For some participants short presentation times may be sufficient to establish a conscious percept, for others the same presentation time might not suffice to allow conscious report of presented scene details. The efficiency of processing briefly flashed visual scenes might therefore not only influence the establishment of the initial scene representation, but may also determine the ability to consciously perceive and report differences between such scenes, for example, whether a kitchen scene was filled with a number of individual kitchen objects or whether the same kitchen was shown empty.

Thus, in order to investigate the influence of individual differences in rapid scene processing on subsequent search, we conducted two experiments. In Experiment 1, participants had to search for predefined target objects embedded in naturalistic scenes (see Castelhana & Henderson, 2007), while we varied the information provided in the flashed previews of the scenes. Each participant was then tested regarding the reportability of preview differences with a post hoc questionnaire, which probed whether the participants had noticed differences between the previews and, if so, in which details the previews differed. This allowed us to divide participants into two groups regarding their reported ability to differentiate between briefly flashed scene preview conditions: The conscious report group

consisted of participants who had processed the previews to such a degree that allowed the report of preview differences, whereas participants in the no report group were unable to report differences among scene previews. We hypothesized that the reported ability to discriminate between flashed scenes might not only indicate the degree of initial processing but could also explain differences in subsequent attention allocation and eye movement control, since these heavily depend on the initial scene representation. In Experiment 2, we further investigated the source of individual differences. Therefore, we retested a subset of participants who had taken part in Experiment 1 using a whole-report task, which allowed assessing TVA parameters regarding the individual processing efficiency of each participant. We assumed that the conscious report group would show higher processing efficiency than the no report group with regard to the TVA parameters visual perceptual processing speed C and VSTM storage capacity K .

EXPERIMENT 1

Experiment 1 investigated the specific contributions of both local and global processing of scene properties during the initial glimpse of a naturalistic scene to the control of subsequent eye movements during visual search. We therefore used the Flash-Preview Moving-Window Paradigm introduced by Castelano and Henderson (2007) to replicate findings of search benefits following identical scene previews. In addition to an identical preview, we modulated the information available during a short scene preview, which either allowed for global processing of the scene or not, i.e., we presented three different preview variations of the same search scene plus a mask as control condition (see Figure 1). These three preview conditions varied in the information available for the participants when being flashed before the search scene: The *background* preview—e.g., an empty kitchen—contained spatial layout information and allowed for scene categorization, whereas the *objects* preview—e.g., a display of typical objects found in a kitchen—lacked spatial layout and could only convey the scene's category indirectly by the need to first identify most objects and then form a category from them. Combining background and objects condition results in the *identical* preview—i.e., a fully equipped kitchen. Thus, the identical and the background condition allow for global processing, whereas the objects preview would—in terms of the contextual guidance model—mainly be processed along the local pathway. We therefore hypothesized that both identical and background previews of the search scene would lead to search benefits, whereas previewing only the objects of the search scene would not benefit subsequent search.

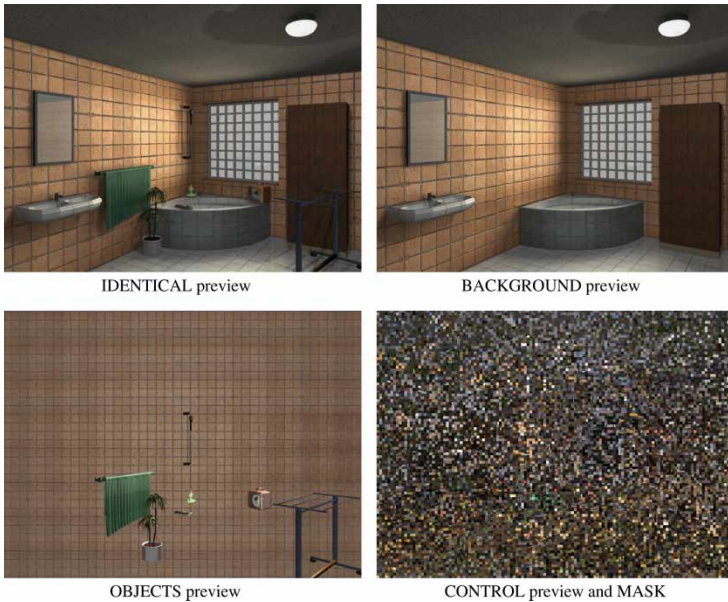


Figure 1. Sample scene previews of Experiment 1 with three different previews of the same bathroom scene (identical, objects, background) and the meaningless control preview also used as a mask. To view this figure in colour, please see the online issue of the Journal.

Additionally, we were interested in whether the two groups of participants differing in their ability to discriminate between preview conditions would also show differences in eye movement behaviour during target search. If the conscious report group is faster at processing visual information and can therefore process briefly presented scene information to a greater degree, this group should show superior target search performance as compared to the no report group. Further, we hypothesized that the differences between the conscious report group and the no report group regarding preview processing would also lead to differential search benefits as a function of information provided in the flashed previews. For instance, the conscious report group should benefit from previews which contain a high degree of information, whereas the no report group should not show such a benefit due to the inability to completely process all the information only briefly provided in the previews.

Method

Participants. Forty students (26 female) from the LMU Munich ranging in age between 19 and 31 ($M = 22.87$, $SD = 2.72$) participated in the study

for course credit or for 8€/hour. All participants reported normal or corrected-to-normal vision and were unfamiliar with the stimulus material.

Stimulus material. The search scenes consisted of 20 3-D-rendered images of real-world scenes. The scenes were displayed on a 19-inch computer screen (resolution 1024×768 pixel, 100 Hz) subtending visual angles of 28.98 degrees (horizontal) and 27.65 degrees (vertical) at a viewing distance of 70 cm. The default background colour was grey (RGB: 51, 51, 51). Each search scene was preceded by either an identical, a background, an objects, or a control preview (see Figure 1 for preview examples) none of which contained the search target. The identical preview was a copy of the search scene except for the missing target object. The background preview resembled the search scene in displaying the same background, but all distinct objects placed on background furnishings were deleted, i.e., the only objects left in the background preview were those that would provide potential support for possible target objects. On the contrary, the objects condition consisted only of distinct objects placed at exactly the same location as in the identical preview but lacking its background. The control was created from scrambled quadratic sections (8×8 pixels) taken from all search scenes and also served as a mask. Thus, the control was meaningless, but contained colours, orientations, and contours as is the case in unscrambled scenes (see Figure 1). Each participant saw each search scene only once, and the four preview conditions for each scene were rotated across participants.

Apparatus. Eye movements were recorded with an EyeLink1000 tower system (SR Research, Canada), which tracks with a resolution of 0.01° visual angle at a sampling rate of 1000 Hz. The position of the right eye was tracked while viewing was binocular. Experimental sessions were carried out on an IBM-compatible display computer running on Windows XP. Stimulus presentation and reaction recording was controlled by Experimental Builder (SR, Research, Canada). The eyetracker was hosted by another IBM-compatible computer running on DOS, which recorded all eye movement data.

Procedure. The procedure of the study phase closely followed the procedure of the “Flash Preview Moving Window” paradigm used in the experiments of Castelhana and Henderson (2007). Experimental sessions were conducted in a moderately lit room (background luminance about 500 lx), in which the illumination was held constant. Each participant received written instructions before being seated in front of the presentation screen. Participants were informed that they would be presented with a series of scenes in which they had to search for a target as fast as possible. They were

also informed that short previews of the scene would precede the display of the search scene and that they should attend to these previews since they could provide additional information.

At the beginning of the experiment, the eyetracker was calibrated for each participant. Therefore, the participants' viewing position was fixed with a chin and forehead rest, followed by a 9-point calibration and validation.

As can be seen in Figure 2, each trial sequence was preceded by a fixation check, i.e., in order to initiate the next trial, the participants had to fixate a cross centred on the screen for 200 ms. When the fixation check was deemed successful, the fixation cross was replaced by the presentation of the scene's preview for 250 ms. After the presentation of a mask for 50 ms, a black target word was displayed at the centre of the grey screen for 2000 ms, which indicated the identity of the target object. Afterwards the search scene was shown through a 2° diameter circular window moving contingent on the participants' fixation location. The rest of the display screen was masked in grey. Thus, no peripheral vision was possible throughout the entire visual search. Participants had to search the scene for the target object and indicate the detection of the target object by holding fixation on the object and pressing a response button. The search scene was displayed for 15 s or until a buttonpress. Three practice trials at the beginning of the experiment allowed participants to get accustomed to the experimental set-up and the restricted vision during search due to the gaze contingent window. At the end of the

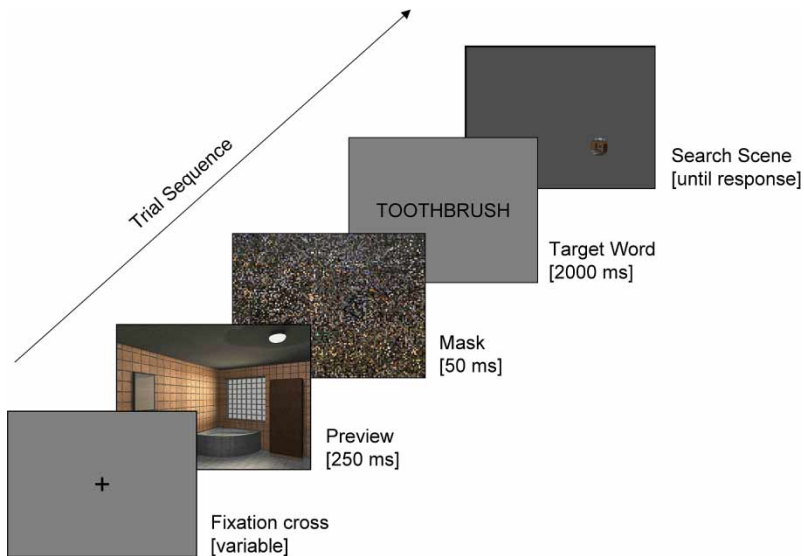


Figure 2. Trial sequence of the “Flash-Preview Moving-Window” paradigm used in Experiment 1. To view this figure in colour, please see the online issue of the Journal.

study phase, participants were asked to fill out a post hoc questionnaire to ascertain whether they were able to distinguish the previews that had been presented. The study phase lasted for about 20 minutes.

Data reduction and statistical analysis. Similar to Castelhamo and Henderson (2007), we analysed a set of behavioural and eye movement data. Response times were calculated from search scene onset until response buttonpress. Error rate was defined as the percentage of those trials, in which participants failed to find and fixate the target object while pressing the response button. Latency to first target fixation was measured from scene onset until the first fixation of the target object. Number of fixations to first target fixation was measured as the sum of all fixations from search scene onset until the first fixation on the target object. Finally, the scan path ratio was defined as the length of the scan pattern, i.e., the sum of all saccade amplitudes until the first fixation of the target object, divided by the shortest distance from the fixation cross to the centre of the target object.

For the analyses of both response time (RT) and eye movement data only correct responses were included, i.e., when the participant pressed the response button while fixating the target object. Additionally, we excluded trials with a fixation number greater than 50 mostly caused by unstable calibration of the gaze contingent window (9.49%).

Further, we had to exclude two participants who showed substantial instabilities in controlling the gaze-contingent window during search. The remaining 38 participants (25 female) ranged in age between 19 and 31 ($M = 22.90$, $SD = 2.75$). After completing Experiment 1, each participant was asked to fill out a questionnaire, which included the question, whether they were able to distinguish the different preview conditions. If a participant claimed to have noticed the preview differences he or she was asked to then describe the different conditions to the instructor in more detail. Only when the participants were able to differentiate between all three scene previews were they assigned to the conscious report group (27 participants). All participants who only noticed a difference between the control preview and “other scenes” were assigned to the no report group (11 participants). The conscious report group (16 female) ranged in age between 19 and 31 ($M = 22.73$, $SD = 2.79$), and the no report group (9 female) ranged in age between 20 and 28 ($M = 23.27$, $SD = 2.87$).

All data were submitted to an analysis of variance (ANOVA) with preview conditions (identical, background, objects, control) as within-subject factor and reportability (conscious report group vs. no report group) as between-subject factor. We confined post hoc tests solely to theoretically driven comparisons of identical versus control, background versus control, objects versus control, and identical versus background preview conditions. Since we expected to find similar patterns across all dependent variables, these

planned contrasts were calculated for all dependent variables and for each participant group.

P-values are reported with exact probabilities when a strong tendency is observed, but fails to reach statistical significance or when *p*-values are presented in tables. All other *p*-values that failed to reach significance are not reported. Further, for effects with multiple degrees of freedom, *p*-values were Greenhouse-Geisser adjusted.

Results

Reaction time data. RT data did not vary significantly across preview conditions, $F(3, 37) = 2.16$. However, both the between-subject factor reportability as well as the interaction of both factors reached significance, $F(3, 37) = 7.36$, $p < .01$, and $F(3, 37) = 4.03$, $p < .05$, respectively (see Figure 3).

Overall, participants who were able to distinguish the three preview conditions showed faster RTs than participants who were not able to distinguish between previews ($M = 4762.92$, $SE = 175.04$ vs. $M = 5574.41$, $SE = 182.18$). Further, preview conditions showed differential preview benefits as a function of reportability (see Table 1). Whereas the conscious report group showed a preview benefit for background versus identical, a tendency for background versus control, and no statistical significance for identical versus control and object versus control, the no report group showed a preview benefit for identical versus control, a tendency for background versus control, and no statistical significance for objects versus control and identical versus background.

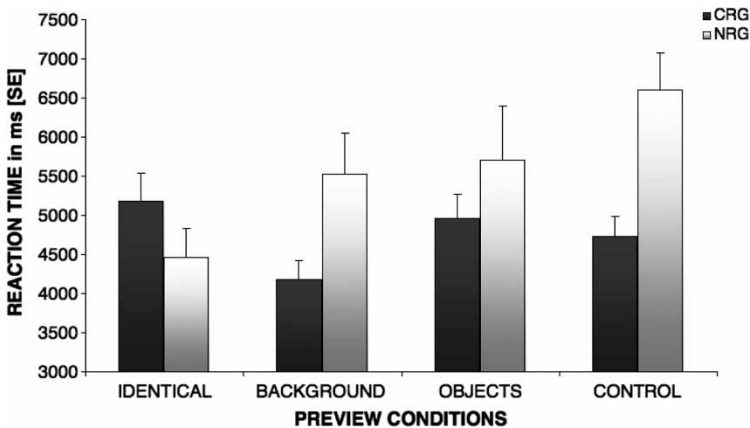


Figure 3. Mean reaction times (standard errors) for visual search in Experiment 1 across preview conditions (identical, background, objects, control) split for participant groups (CRG = conscious report group, NRG = no report group).

TABLE 1

Summary of planned contrasts for response time and latency to first target fixation across preview conditions during visual search in Experiment 1 split for participant groups

| <i>Planned contrasts</i> | | <i>t</i> | <i>df</i> | <i>p</i> |
|----------------------------------|---------|----------|-----------|----------|
| Response time | | | | |
| CRG | I vs. C | 1.00 | 26 | .16 |
| | B vs. C | 1.63 | 26 | .05 |
| | O vs. C | 0.58 | 26 | .28 |
| | I vs. B | 2.72 | 26 | .01 |
| NRG | I vs. C | 4.00 | 10 | .00 |
| | B vs. C | 1.59 | 10 | .07 |
| | O vs. C | 1.05 | 10 | .16 |
| | I vs. B | 1.47 | 10 | .09 |
| Latency to first target fixation | | | | |
| All | I vs. C | 1.27 | 37 | .11 |
| | B vs. C | 2.66 | 37 | .01 |
| | O vs. C | 1.12 | 37 | .13 |
| | I vs. B | 0.65 | 37 | .26 |
| CRG | I vs. C | 0.51 | 26 | .31 |
| | B vs. C | 2.00 | 26 | .03 |
| | O vs. C | 0.11 | 26 | .46 |
| | I vs. B | 2.13 | 26 | .02 |
| NRG | I vs. C | 3.85 | 10 | .00 |
| | B vs. C | 1.76 | 10 | .05 |
| | O vs. C | 1.63 | 10 | .07 |
| | I vs. B | 1.34 | 10 | .11 |

I = identical, B = background, O = objects, C = control; CRG = conscious report group, NRG = no report group.

Error rate. Error rates averaged at 21.32% (identical: 18.11%; background: 24.71%; objects: 14.66%; control: 27.78%). There was a nearly significant main effect of preview, $F(3, 37) = 2.66$, $p = .05$, but no effect of group, $F(3, 37) < 1$, and no interaction, $F(3, 37) = 1.09$. Participants produced more errors during target search when they had been presented with a control preview than with an identical or an objects preview, $t(37) = 2.35$, $p < .05$. All other contrasts failed to reach significance.

Latency to first target fixation. For latency data, all main effects as well as their interaction were significant. There was a main effect of preview, $F(3, 37) = 3.14$, $p < .05$, a main effect of reportability, $F(3, 37) = 11.15$, $p < .01$, and a significant interaction, $F(3, 37) = 4.24$, $p < .01$ (see Figure 4).

As can be seen in Table 1, planned contrasts revealed a significant preview benefit for background versus control across all participants; all other planned contrasts failed to reach significance. Overall, participants who were

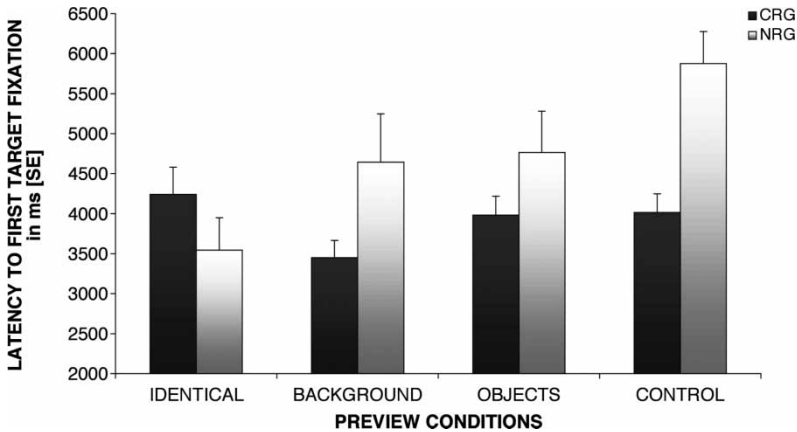


Figure 4. Mean latencies (standard errors) for visual search in Experiment 1 across preview conditions (identical, background, objects, control) split for participant groups (CRG =conscious report group, NRG =no report group).

able to distinguish the three preview conditions showed shorter latencies than participants who were not able to distinguish between previews ($M = 3921.56$, $SE = 142.49$ vs. $M = 4707.82$, $SE = 110.82$). The interaction of the factors preview and reportability can be characterized as the following: The conscious report group showed strong search benefits after the presentation of the background preview compared to control as well as to the identical preview. All other planned contrasts failed to reach significance. On the other hand, the no report group showed a graded effect of preview conditions in that the effect was strongest for the Identical preview and showed a strong tendency to decrease for objects and background conditions.

Thus, the main effect of preview was caused by a significant search benefit following a background preview as compared to the presentation of the control and preview benefit significantly varied as a function of groups.

Number of fixations until target fixation. There was a significant interaction between preview conditions and the reportability factor, $F(3, 37) = 4.48$, $p < .01$, and the main effects of preview and reportability showed trends, $F(3, 37) = 2.44$, $p = .07$, and $F(3, 37) = 3.18$, $p = .08$, respectively (see Figure 5).

As can be seen in Table 2, planned contrasts for the preview conditions showed a search benefit for background versus control; none of the other contrasts reached significance. For the grouping factor, the conscious report group showed a trend to need fewer fixations until the first target fixation than no report group ($M = 13.39$, $SE = 0.45$ vs. $M = 14.82$, $SE = 0.58$). The significant interaction between both factors for the number of fixations

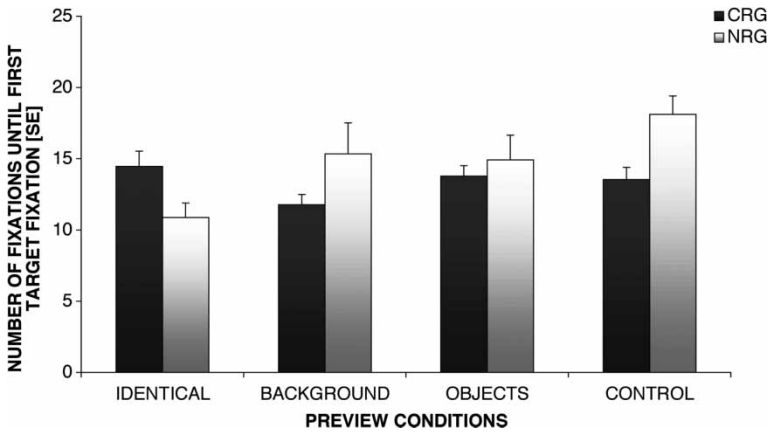


Figure 5. Mean number of fixations until first target fixation (standard errors) for visual search in Experiment 1 across preview conditions (identical, background, objects, control) split for participant groups (CRG =conscious report group, NRG =no report group).

largely resembles the interaction observed for the latency to first target fixation: Again, the conscious report group showed strong search benefits after the presentation of the background preview compared to control as well as to the identical preview. The other contrasts failed to reach significance. However, the no report group showed a strong search benefit following an identical preview and a tendency for fewer fixations for Identical as compared to the background condition. There were no

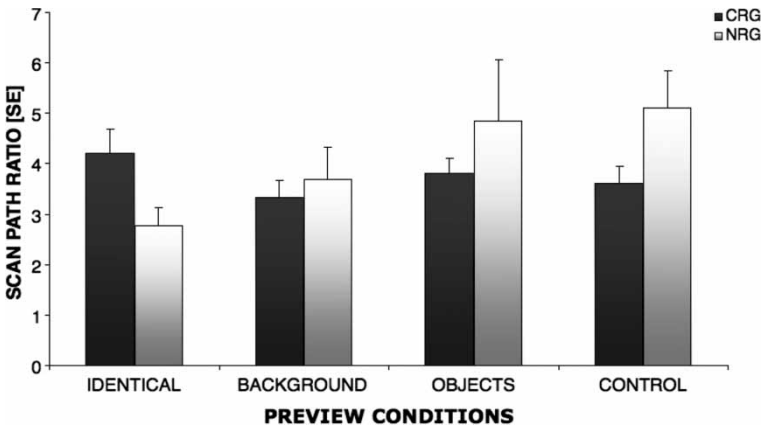


Figure 6. Mean scan path ratio (standard errors) for visual search in Experiment 1 across preview conditions (identical, background, objects, control) split for participant groups (CRG =conscious report group, NRG =no report group).

TABLE 2

Summary of planned contrasts for number of fixations until target fixation and scan path ratio across preview conditions during visual search in Experiment 1 split for participant groups

| <i>Planned contrasts</i> | | <i>t</i> | <i>df</i> | <i>p</i> |
|---|---------|----------|-----------|----------|
| Number of fixations until target fixation | | | | |
| All | I vs. C | 1.13 | 37 | .13 |
| | B vs. C | 1.21 | 37 | .02 |
| | O vs. C | 0.74 | 37 | .23 |
| | I vs. B | 0.49 | 37 | .31 |
| CRG | I vs. C | 0.65 | 26 | .26 |
| | B vs. C | 1.74 | 26 | .05 |
| | O vs. C | 0.11 | 26 | .42 |
| | I vs. B | 2.13 | 26 | .02 |
| NRG | I vs. C | 3.83 | 10 | .00 |
| | B vs. C | 1.32 | 10 | .11 |
| | O vs. C | 1.44 | 10 | .08 |
| | I vs. B | 1.72 | 10 | .06 |
| Scan path ratio | | | | |
| CRG | I vs. C | 0.97 | 26 | .17 |
| | B vs. C | 0.50 | 26 | .31 |
| | O vs. C | 0.45 | 26 | .33 |
| | I vs. B | 1.52 | 26 | .07 |
| NRG | I vs. C | 3.30 | 10 | .00 |
| | B vs. C | 1.38 | 10 | .10 |
| | O vs. C | 0.19 | 10 | .43 |
| | I vs. B | 1.26 | 10 | .12 |

I = identical, B = background, O = objects, C = control; CRG = conscious report group, NRG = no report group.

significant differences between background versus control and objects versus control.

Scan path ratio. The ANOVA for the scan path ratio did not show significant main effects of preview or reportability, $F(3, 37) = 1.51$, and $F(3, 37) = 1.28$, respectively. However, there was a strong tendency for a significant interaction of both factors, $F(3, 37) = 2.62$, $p = .05$ (see Figure 6).

As can be seen in Table 2, no significant differences were found across preview conditions for the conscious report group. There was a tendency for a smaller scan path ratio for identical as compared to background preview. However, planned contrasts for the no report group revealed a significantly decreased scan path ratio for identical versus control preview.

Thus, all dependent variables except for error rates showed that the effects for preview conditions strongly interacted with the ability to differentiate between the three preview conditions. Participants who reported not being

able to differentiate between previews showed greatest search benefit for identical previews, whereas the group who did report being able to differentiate between previews needed the least number of fixations to find the target in the background preview condition. For this group of participants the identical preview did not lead to search benefits compared to the control condition.

Discussion

Experiment 1 investigated the influence of both local and global processing of scene properties during the initial glimpse of a naturalistic scene on the control of subsequent eye movements during visual search. We were also interested in whether the ability to distinguish between flashed scenes would modulate search performance. The results reported here allow a number of interesting implications:

First, we found strong evidence for the influence of flashed scene previews on the guidance of eye movements which indicates that participants were able to generate, store, and make use of an initial scene representation for subsequent target search. Since the visual field was restricted to a 2° diameter gaze-contingent window, the search benefits can be attributed to the information gathered from the initial glimpse of the scene. The initial scene representation must then be stored across saccades in order to play a functional role in eye movement control. This is in line with findings of Castelano and Henderson (2007), who were able to show that the scene representation, which is used to effectively guide attention through search space is robust to a 2 s delay between preview presentation and search scene display as well as to several intervening saccades. Thus, the results of Experiment 1 add to the growing evidence that initially generated scene representations can be stored across saccades continuously exhibiting their influence on eye movement control.

Second, compared to the control condition, participants were faster and needed fewer fixations to find and fixate the target object when being presented only with the scene's background. This shows that a scene representation generated from a preview, which mainly provides global scene information and therefore allows rapid setup of scene priors, can already lead to an effective deployment of attention during subsequent search.

Third, the conscious report group was overall faster (RT and latency to first target fixation) and needed fewer fixations to find the target as compared to the no report group. It seems that participants who had processed the previews to a greater degree were also able to control their eye movements more effectively during search. Additionally, the interaction found across all dependent variables showed that the two groups of

participants, which differed in their reported ability to process the previews, also differed in their ability to benefit from scene previews. The main characteristic of the interaction, which was observable across RT, latency to first target fixation, and fixation data was that although the conscious report group showed the greatest search benefits after the presentation of background previews, the no report group profited most from identical scene previews. This modulation of preview benefits as a function of group led us to imply that the locus of the individual differences observed in search performance might lie in varying degrees of preview processing.

EXPERIMENT 2

Experiment 2 aimed at further investigating the locus of effects for the individual differences observed during visual search in Experiment 1. We therefore retested participants that had previously taken part in Experiment 1 using a TVA based whole-report task, which provided us with information on the general processing efficiency of each participant. This way we were able to examine whether the two participant groups differed in perceptual processing speed and/or VSTM storage capacity.

Methods

Participants. Twenty-five students (18 female) from the LMU Munich ranging in age between 19 and 28 ($M = 23.24$, $SD = 2.55$) participated in the whole-report task for course credit or for 8€/hour. All participants reported normal or corrected-to-normal vision and were unfamiliar with the stimulus material. Further, all 25 participants had taken part in Experiment 1, 11 of whom had been previously assigned to the no report group and 14 to the conscious report group.

Stimulus material. For the whole-report task, five red target letters (each 0.5° high \times 0.4° wide) were presented in a vertical column, 2.5° of visual angle either to the left or to the right of a fixation cross, on a black screen. Stimuli for a given trial were randomly chosen from a prespecified set of letters (ABEFHJKLMNPRSTWXYZ), with the same letter appearing only once per trial. In some trials letter displays were masked. Masks consisted of letter-sized squares (of 0.5°) filled with a “+” and an “×”.

Apparatus. The TVA experiment was conducted in a dimly lit, sound-proof cubicle. Stimuli were presented on a 17-inch monitor (1024×768 pixel screen resolution, 70 Hz refresh rate). Subjects viewed the monitor from a distance of 50 cm, controlled by the aid of a head- and chinrest.

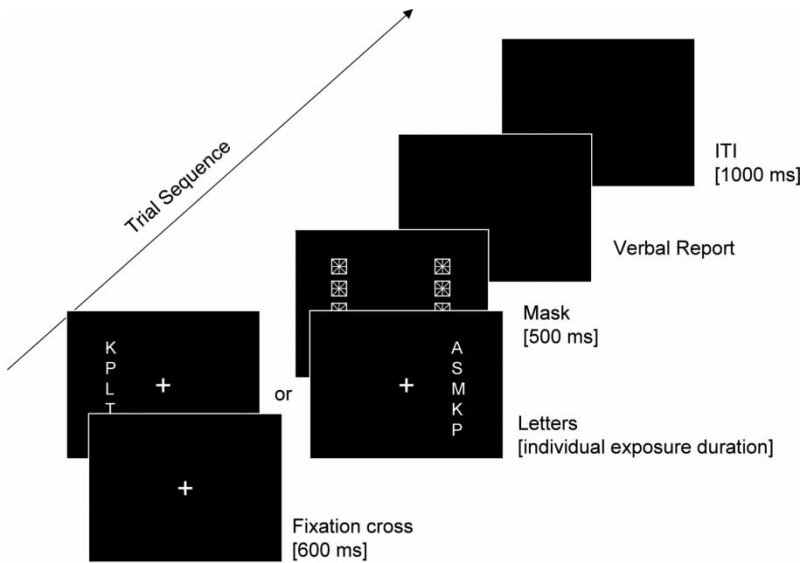


Figure 7. Trial sequence of the whole-report task used in Experiment 2.

Procedure. Figure 7 shows the trial sequence of the whole-report task. Participants were first instructed to fixate a white cross ($0.3^\circ \times 0.3^\circ$) presented for 600 ms in the centre of the screen on a black background. Then five red target letters were presented in a vertical column either to the left or to the right of the fixation cross. The participants had to report as many letters as possible. The experiment comprised two phases: In Phase 1 (pretest), three exposure durations of the target letters were determined for Phase 2 (main test), in which the data were collected. The pretest comprised 24 masked trials with an exposure duration of 86 ms. It was assessed whether the subject could, on average, report one letter (20%) per trial correctly. If this was achieved, exposure durations of 43 ms, 86 ms, and 157 ms were used in the main test. Otherwise, longer exposure durations of 86 ms, 157 ms, and 300 ms were used. Here, letter displays were presented either masked or unmasked. The masks were presented for 500 ms at each letter location. Due to “iconic-memory” buffering, the effective exposure durations are usually prolonged by several hundred milliseconds in unmasked as compared to masked conditions (Sperling, 1960). Thus, by factorially combining the three exposure durations with the two masking conditions, six different “effective” exposure durations were produced. These were expected to generate a broad range of performance, tracking the early and the late parts of the functions relating response accuracy to effective exposure duration. In several previous studies that used a similar paradigm (e.g., Finke, Bublak, Krummenacher, Kyllingsbaek, Müller, & Schneider, 2005), highly reliable estimates of the

parameters C and K were obtained on the basis of 16 trials per target condition. On this basis, each subject completed 288 trials (2 hemifields \times 2 masking conditions \times 3 exposure durations \times 16 trials per target condition) in the present experiment. Before each phase, subjects were given written and verbal instructions.

Results

The experimental results of the whole-report task are described by the TVA parameter estimates for “visual perceptual processing speed” and “VSTM storage capacity”. These parameters were estimated using the standard procedure introduced by Duncan et al. (1999) and used in several other recent studies (e.g., Bublak et al., 2005; Finke, Bublak, Dose, Müller, & Schneider, 2006; Habekost & Rostrup, 2007; Hung, Driver, & Walsh, 2005). Detailed descriptions on the software used can be found in Kyllingsbaek (2006), and detailed neural interpretations of the mathematically specified TVA concepts are described in Bundesen, Habekost, and Kyllingsbaek (2005). In short, the probability of identifying a given object x is modelled by an exponential growth function. The slope of this function indicates the total rate of information uptake in objects per second (perceptual processing speed, denoted by C), and its asymptote the maximum number of objects that can be represented at a time in VSTM (VSTM storage capacity, K).

Since we were interested in whether the difference in reportability of preview differences was due to the participants’ processing efficiency, we compared the 14 participants of the conscious report group with 11 participants of the no report group regarding both the perceptual processing speed C and the VSTM storage capacity K .

Perceptual processing speed C . C is defined as a measure of the perceptual processing speed in elements/second. C across all participants ranged from 6.22 to 33.46 ($M = 17.06$, $SD = 7.10$). Planned contrasts showed that participants from the conscious report group ($M = 19.79$, $SE = 1.74$) are characterized by a higher perceptual processing speed C than participants from the no report group ($M = 13.60$, $SE = 1.96$), $t(1) = 2.36$, $p = .01$.

VSTM storage capacity K . Parameter K reflects the number of items that can be simultaneously maintained in VSTM. K across all participants ranged from 2.38 to 4.00 ($M = 3.31$, $SD = 0.57$). There was no significant difference between groups regarding VSTM storage capacity K (conscious report group: $M = 3.43$, $SE = 0.15$ vs. no report group: $M = 3.17$, $SE = 0.17$), $t(1) = 1.11$, $p > .05$.

Thus, the TVA parameters show that although participants of the conscious report group did not differ from the no report group in terms of VSTM storage capacity, they seem to have been able to process information faster than the no report group.

Discussion

Experiment 2 was set out to further investigate the nature of individual differences observed in Experiment 1, where participants showed differences in their efficiency to search for target objects in naturalistic scenes. We had hypothesized that the differences in the reportability of preview differences were due to varying degrees of information processing efficiency across participants leading to differential effects on eye movement control. Since post hoc questionnaires can only provide subjective measures of information processing, we conducted a follow-up experiment on the basis of the TVA, i.e., a whole-report task using simple letters as stimulus material, which has shown to provide reliable estimates of individual processing efficiency parameters (e.g., Bublak et al., 2005; Finke et al., 2006; Habekost & Rostrup, 2007; Hung et al., 2005). In this experiment, we observed a higher perceptual processing speed C for the conscious report group than for the no report group, and neither group differed in VSTM storage capacity K . These findings shed more light on the locus of the individual differences that emerge when being presented with only briefly visible scene previews: It seems that those participants who were able to distinguish between different scene previews were able to do so due to a higher degree of processing speed.

According to the TVA model, which is strongly related to the biased-competition conceptualization of visual attention (Desimone & Duncan, 1995), visual objects are processed in parallel and compete for selection (i.e., conscious representation). In TVA, selection of an object is synonymous with its encoding into limited-capacity VSTM, i.e., its “conscious” representation within the information processing system. Objects that are selected and hence may be reported from a briefly exposed visual display are those elements for which the encoding is completed before the sensory representation of the stimulus array has decayed and before VSTM has filled up with other objects. Thus, when visual input is only available for a very limited amount of time, the number of items that can be encoded into VSTM greatly depends on the speed of processing visual information. Even though the whole-report task of Experiment 2 used much simpler stimulus material than the scenes presented during Experiment 1, it seems that the higher processing speed observed for processing letters enabled participants from the conscious report group to better distinguish between the briefly flashed scene previews by extracting more detailed information than the no report

group. VSTM storage capacity, on the other hand, did not seem to play a decisive role in distinguishing scene previews and effectively controlling eye movements during subsequent search. As we will discuss in further detail, a higher processing speed while leading to increased performance in the TVA whole-report task, might not only benefit when searching for target objects in naturalistic scenes.

GENERAL DISCUSSION

One goal of the present set of experiments was to examine the contribution of both global and local processing to the initial scene representation, which can be rapidly established from a first glimpse of a complex scene. We were interested in how this initially crude visual representation can control the deployment of attention and eye movements during subsequent target search, while more detailed object information is continuously added to the evolving scene representation. Additionally we further investigated the role that individual differences play in the generation of initial scene representations and how these can modulate eye movement behaviour during target search. The results of this study show that the consideration of individual differences in information processing efficiency allows a more detailed understanding of the cognitive processes that underlie the processing of visual scene information and subsequent eye movement control.

Dominance of global processing during the first glimpse of a scene

In Experiment 1, we varied the information provided during flashed previews of the search scenes in order to investigate the influence of both local and global processing on the control of subsequent eye movements during visual search. According to the contextual guidance model (Torralba et al., 2006), target detection is achieved by estimating the probability of the presence of the target object at different locations given the combined output of both local and global processing and moving the eyes to the location with the highest target probability. However, before attention is located to a particular part of a scene, scene context activates scene priors, which then allow the restriction of search space to those locations that are most probable to contain the target. The observed search benefit following a preview that only contained a scene's background implies that processing global features to compute spatial layout and set up scene priors combined with task knowledge seems to suffice to restrict search to highly probable locations in a scene. On the other hand, when the preview only contained individual objects, but lacked spatial layout and the possibility to quickly set

up scene priors, there was no observable preview benefit indicating that the local processing of individual objects in a scene is not as beneficial. In terms of the contextual guidance model, this suggests that isolated objects without being embedded in a broader scene context do not allow for enough contextual guidance to effectively control eye movements, whereas processing along the global pathway does allow effective eye movement control without the necessity to additionally segregate and compute all displayed objects. This is in line with prior work that has shown that the computation of a scene's gist can be done very rapidly (e.g., Oliva & Schyns, 1997; Oliva & Torralba, 2006; Potter, 1975; Thorpe, Fize, & Marlot, 1996), whereas only a few objects can be identified within a split second thus preventing the establishment of a complete mental representation of a scene with all identities and visual details of objects within the first glimpse (e.g., Castelhamo & Henderson, 2005; Henderson & Hollingworth, 2003; Tatler et al., 2003). For example, the gist of a scene can be inferred from its spatial layout, its global scene properties, or simply the spatial distribution of colours, major scales, and orientations (e.g., Greene & Oliva, 2006; McCotter, Gosselin, Sowden, & Schyns; Oliva & Schyns, 2000; Schyns & Oliva, 1994). Accordingly, accuracy in scene recognition is not affected by the quantity of objects in a scene and can be achieved equally well when local object recognition is hampered by blur (Oliva & Schyns, 1997; Schyns & Oliva, 1994; for a review see Oliva, 2005).

Thus, although it is not possible to fully separate the effects of local and global scene processing and more work on this topic is needed, we argue from our data that global scene processing is a prerequisite for the rapid generation of an initial scene representation that allows effectively control of subsequent eye movements; local object processing is not.

A glimpse is not a glimpse

In Experiment 1, we found a main effect of the between-subject factor on reaction times and latency to first target fixation and a strong trend for the number of fixations in that the conscious report group generally showed superior search performance as compared to the no report group. What are the underlying cognitive processes that cause these observable differences? Verbal reports of participants may only provide subjective and indirect information and have to be treated with reserve. However, the reported inability to differentiate between the three preview conditions does imply a reduced degree of processing during a flashed preview as compared to the conscious report group. When presented with a flashed preview of a scene, it is likely that the first wave of feedforward processing in the visual brain is followed by a series of more complex processes required for generating

conscious perception of the scene (Kirchner & Thorpe, 2006; for a review see Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). Given the limited and brief presentation of a preview, for some subjects short presentation times may be sufficient to establish a conscious percept, for others the same presentation time might not suffice to allow conscious report of presented scene details. The efficiency of processing briefly flashed visual scenes might therefore not only influence the establishment of the initial scene representation, but may also determine the ability to consciously perceive and report differences between such scenes, for example, whether a kitchen scene was filled with a number of individual kitchen objects or whether the same kitchen was shown empty.

The hypothesis that group differences in preview reportability were due to varying degrees of processing within the first glimpse of a scene was further supported by Experiment 2, which provided evidence that the conscious report group is able to process briefly flashed visual information faster than the no report group. Thus, it seems that the ability to efficiently process simple letters might also enable the extraction of more detailed information from scene previews, which can be used subsequently to efficiently control eye movements in the search for a predefined target object.

Interaction of global and local pathways as a function of the degree of processing

At first glance, it seems surprising that contrary to the sparser background preview the identical preview did not result in significant search benefits. Since the identical and background previews share the same global features, restriction of search space by a combination of setting up scene priors, spatial layout computation, and task knowledge should be possible for both previews alike. Also, Castelhana and Henderson (2007) observed clear search benefits when presenting identical as compared to the meaningless previews even when these were downscaled in size. However, taking individual differences in preview processing into account might be able to explain these seemingly contradictory findings.

Both experiments taken together provided clear evidence for a strong interaction between the degree of preview processing and the degree of information available in the different preview conditions. Participants who had reported being able to distinguish between the three flashed preview conditions (Experiment 1) and who showed a greater perceptual processing speed (Experiment 2) benefited most from the background preview of the search scene, whereas participants who had reported not being able to distinguish between the three preview conditions and who were characterized

by a lower processing speed searched the scenes most efficiently when presented with an identical preview.

Thus, contrary to our hypothesis, the no report group benefited most from the identical preview despite their slower processing speed. It seems that although this participant group was not able to report differences between scene previews, the varying degree of scene information was nevertheless retained and subsequently affected search efficiency. This could be due to implicit priming effects, which would lead to a higher activation of a scene category the more information is contained in a preview. Since the no report group seems to be unable to process individual objects embedded in a flashed scene preview to a reportable degree, these additional objects in the identical preview condition might have contributed to the activation of a scene category in a unambiguous way.

On the other hand, the conscious report group—which represents about two-thirds of all participants—did not significantly benefit from an identical preview as compared to the control condition although it contained more information than the background preview. It seems as if the additional objects in the identical preview led to detrimental effects on the generation of an initial scene representation when the flashed scene was processed to a higher degree. In the contextual guidance model (Torralba et al., 2006), the setup of scene priors solely takes place on the global pathway, which parallels the processing of local objects; the output of both pathways is later combined to interact in a scene-modulated saliency map which controls eye movement behaviour. The background preview provides less but unequivocal information needed for the setup of scene priors due to its predominant processing on the global pathway, and the identical preview additionally provides local object information, which can be processed parallel to the global pathway before their outputs combine. We argue that due to the enhanced processing speed of the conscious report group more objects can be segmented from the background and processed up to the level of identification, which in turn might activate additional priors regarding scenes or objects generated along the local pathway, i.e., plates and glasses on a dining room table could also elicit scene priors related to the context “kitchen”. In line with these considerations, there is evidence that over the course of time contingencies between objects are learned such that the perception of one object can generate strong expectations about the probable presence and location of other objects (e.g., Chun & Jiang, 1999; Green & Hummel, 2006; for a review see Oliva & Torralba, 2007). Thus, the conscious report group, in contrast to the no report group, might activate competing scene priors, i.e., one generated along the global pathway and another generated along the local pathway. This competition needs to be resolved leading to detrimental effects on the effective control of subsequent search behaviour. The no report group, on the other hand, cannot process local

information to such an extent that individual objects presented in the identical preview could elicit locally generated scene priors. In this case no detrimental competition amongst equivocal scene categories impedes effective eye movement control.

How strong object and background information can interfere was demonstrated in a study by Joubert, Rousselet, Fize, and Fabre-Thorpe (2007) where they showed that the processing of scene context is fast enough to allow for early interactions between object and context processing. They used a go/no-go rapid visual categorization task in which participants had to distinguish as fast as possible whether a scene that was only present for 26 ms was a “man-made environment” or a “natural environment”. An interesting finding was that the presence of a salient object in a scene delayed processing of the background and induced an accuracy drop of up to 4.8%. When an object was also incongruent with the scene context, its detrimental effects on scene categorization were further increased. Similarly, Davenport and Potter (2004) had found evidence for an early interaction between scene background and objects in that inconsistent objects led to decreased performance in an object and background naming task. Although not intended, some of the objects in the scenes presented in our study might not have led to the same setup of scene priors as the ones generated by a scene background. For participants with a high processing speed, this might result in detrimental background–object interactions when presented with an identical preview containing both background and object information, which in turn could impede the effective restriction of search space thereafter. In line with this reasoning, Lavie and Cox (1997) had shown in a visual search experiment that efficient target search led to greater distractor interference than inefficient search. This seemingly paradoxical finding follows the load hypothesis, according to which easier searches should leave enough spare attentional resources to spill over to the processing of distractors. Applied to our findings, this would imply that the greater processing speed of the conscious report group led to a greater degree of distractor processing and therefore distractor interference. We further argue that distractor interference would be especially detrimental, when there are objects in the scene that could also be diagnostic for a different scene category. Accordingly, the conscious report group would be more susceptible to be misled by such objects than the no report group.

Contrary to Castelhamo and Henderson (2007), we only found search benefits following identical previews for the participants of the no report group, which only accounted for a third of all participants. Thus, we did not observe an overall search benefit for identical previews across all participants. A possible explanation for these contradictory findings could be the manipulation of previews used in our study as compared to the ones used in the Castelhamo and Henderson study. In the latter, identical, different,

concept, or miniature previews were compared across several experiments, whereas we contrasted different versions of the same scene preview, which only varied in the amount of information presented. In particular, the conscious report group might have been distracted by the presence of objects in the identical preview, since they reported noticing the absence or presence of objects across previews. Another reason for the diverging results could be that we used 3-D-rendered scenes, whereas Castelhamo and Henderson used photographs of scenes. Although the 3-D-rendered scenes are very realistic, their scene composition might be more artificial than that of photographs. An artificially created scene generally tends to contain fewer and more isolated objects and might therefore be less cluttered than when simply taking a picture of a real living room. This might have caused the objects displayed in our scenes to be more salient than the ones used by Castelhamo and Henderson, increasing the possibility of detrimental effects especially in the conscious report group.

CONCLUSION

The study presented here provided evidence for the predominance of global processing in the generation of initial scene representations and in the effective control of attention and eye movements during visual search in naturalistic scenes. A short glimpse of a scene's background suffices to restrict search space when subsequently looking for a predefined target object. Additionally, we found that people greatly differ in their ability to process flashed scene previews. We argue that varying degrees of processing either local or global scene information can lead to differential generations of initial scene representations, which could account for the individual differences observed in subsequent eye movement behaviour. These results provide evidence for individual differences in the processing efficiency during a short glimpse of a naturalistic scene, but future work will need to further investigate the relationship between varying degrees of information processing and degrees of effective attention and eye movement control, for example, by actively manipulating the degree of scene processing. Although the approach to the investigation of individual differences in scene processing offered in this study is by no means meant to be exhaustive, we want to bring attention to the fact that individual differences can arise and should not be neglected when investigating the impact of the first glimpse of a scene.

REFERENCES

- Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, 39, 2947–2953.

- Brockmole, J. R., Hambrick, D. Z., Windisch, D. J., & Henderson, J. M. (2008). The role of meaning in contextual cueing: Evidence from chess expertise. *Quarterly Journal of Experimental Psychology*, *61*, 1886–1896.
- Bublak, P., Finke, K., Krummenacher, J., Preger, R., Kyllingsbaek, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention II: Evidence from two patients with frontal or parietal damage. *Journal of the International Neuropsychological Society*, *11*, 843–854.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523–547.
- Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *353*, 1271–1281.
- Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*, 291–328.
- Castelhano, M., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition*, *12*(6), 1017–1040.
- Castelhano, M., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 753–763.
- Castelhano, M. S., & Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology*, *62*(1), 1–14.
- Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences*, *102*(35), 12629–12633.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*, 360–365.
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564.
- Dehaene, S., Changeux, J., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, *128*, 450–478.
- Eden, G. E., Stein, J. E., Wood, H. M., & Wood, E. B. (1994). Difference in eye movements and reading problems in dyslexic and normal children. *Vision Research*, *34*, 1345–1358.
- Finke, K., Bublak, P., Dose, M., Müller, H. J., & Schneider, W. X. (2006). Parameter-based assessment of spatial and non-spatial attentional deficits in Huntington's disease. *Brain*, *129*, 1137–1151.
- Finke, K., Bublak, P., Krummenacher, J., Kyllingsbaek, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention: I. evidence from normal subjects. *Journal of the International Neuropsychological Society*, *11*, 832–842.
- Green, C., & Hummel, J.E. (2006). Familiar interacting object pairs are perceptually grouped. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1107–1119.
- Greene, M. R., & Oliva, A. (2006). Natural scene categorization from the conjunction of ecological global properties. In R. Sun (Ed.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 291–296). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Habekost, T., & Rostrup, E. (2007). Visual attention capacity after right hemisphere lesions. *Neuropsychologia*, *45*(7), 1474–1488.
- Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16*(4), 219–222.

- Henderson, J. M., & Castelano, M. S. (2005). Eye movements and visual memory for scenes. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 213–235). New York: Oxford University Press.
- Henderson, J. M., & Hollingworth, A. (2003). Eye movements, visual memory, and scene representation. In M. A. Peterson & G. Rhodes (Eds.), *Analytic and holistic processes in the perception of faces, objects, and scenes* (pp. 356–383). New York: Oxford University Press.
- Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 396–411.
- Hung, J., Driver, J., & Walsh, V. (2005). Visual selection and posterior parietal cortex: Effects of repetitive transcranial magnetic stimulation on partial report analyzed by Bundesen's theory of visual attention. *Journal of Neuroscience*, *25*(42), 9602–9612.
- Hutzel, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language*, *89*(1), 235–242.
- Itti, L., & Koch, C. (2000). A saliency-based mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 1254–1259.
- Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*, 3286–3297.
- Kirchner, H., & Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*, 1762–1776.
- Kyllingsbaek, S. (2006). Modeling visual attention. *Behavior Research Methods*, *38*(1), 123–133.
- Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science*, *8*(5), 395–398.
- McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. G. (2006). The use of visual information in natural scenes. *Visual Cognition*, *12*, 938–953.
- Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*, 205–231.
- Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, *100*, 11163–11170.
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier.
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176–210.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36.
- Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.
- Potter, M. C. (1975). Meaning in visual scenes. *Science*, *187*, 965–966.
- Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, *47*, 2714–2726.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.

- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74.
- Tatler, B. W., Gilchrist, I. D., & Land, M. F. (2005). Visual memory for objects in natural scenes: From fixations to object files. *Quarterly Journal of Experimental Psychology*, 58A(5), 931–960.
- Tatler, B. W., Gilchrist, I. D., & Rusted, J. (2003). The time course of abstract visual representation. *Perception*, 32, 579–592.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in human visual system. *Nature*, 381, 520–522.
- 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.
- Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, 46, 629–646.
- Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruity influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, 59(11), 1931–1949.
- Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, 18, 321–342.
- Werner, S., & Thies, B. (2000). Is “change blindness” attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, 7(1–3), 163–173.

Manuscript received November 2007

Manuscript accepted October 2008

First published online December 2008