Routledge
Taylor & Francis Group

Eyeing the eyes in social scenes: Evidence for top-down control of stimulus selection in simultanagnosia

Kirsten A. Dalrymple[1], Alexander K. Gray[1], Brielle L. Perler[1], Elina Birmingham[2], Walter F. Bischof[3], Jason J. S. Barton[1,4,5], and Alan Kingstone[1]

[1]Department of Psychology, University of British Columbia, Vancouver, Canada

[2]Faculty of Education, Simon Fraser University, Burnaby, Canada

[3]Department of Computing Science, University of Alberta, Edmonton, Canada

[4]Department of Medicine (Neurology), University of British Columbia, Vancouver, Canada

[5]Department of Ophthalmology & Visual Sciences, University of British Columbia, Vancouver, Canada

Simultanagnosia is a disorder of visual attention resulting from bilateral parieto-occipital lesions. Healthy individuals look at eyes to infer people's attentional states, but simultanagnosics allocate abnormally few fixations to eyes in scenes. It is unclear why simultanagnosics fail to fixate eyes, but it might reflect that they are (a) unable to locate and fixate them, or (b) do not prioritize attentional states. We compared eye movements of simultanagnosic G.B. to those of healthy subjects viewing scenes normally or through a restricted window of vision. They described scenes and explicitly inferred attentional states of people in scenes. G.B. and subjects viewing scenes through a restricted window made few fixations on eyes when describing scenes, yet increased fixations on eyes when inferring attention. Thus G.B. understands that eyes are important for inferring attentional states and can exert top-down control to seek out and process the gaze of others when attentional states are of interest.

Keywords: Attention; Eye gaze; Neuropsychology; Simultanagnosia; Social scenes perception.

Dorsal simultanagnosia is a disorder of visual attention that results in an inability to see more than one object at a time (Holmes & Horrax, 1919; Rafal, 1997). It is typically the consequence of bilateral lesions to the parieto-occipital junction (Bálint, 1909; Holmes & Horrax, 1919; Riddoch et al., 2010; Rizzo & Vecera, 2002) and may

occur in the context of Bálint syndrome, though it can also occur independently (Damasio, 1985). Sometimes these patients can only see pieces of the objects in their visual world, unaware that they are seeing just one component of a larger form. Thus, when viewing single (global) forms made up of smaller (local) elements in hierarchical

stimuli (Navon, 1977), patients report the local elements but not the global aspect, a phenomenon known as "local capture" (Clavagnier, Fruhmann Berger, Klockgether, Moskau, & Karnath, 2006; Dalrymple, Bischof, Cameron, Barton, & Kingstone, 2009; Dalrymple, Kingstone, & Barton, 2007; Karnath, Ferber, Rorden, & Driver, 2000). Notably, their global perception improves when local elements become smaller or more densely packed, emphasizing the modulating impact of stimulus properties on attentional deficits (Clavagnier et al., 2006; Dalrymple et al., 2007).

Dorsal simultanagnosia has been popularly understood as a restriction of object-based visual attention: Patients can only see one object at a time, at the expense of all other objects (Baylis, Driver, Baylis, & Rafal, 1994; Moreaud, 2003; Rafal, 2003). However, others have proposed that simultanagnosia may involve a space-based limitation of attention. For example, some authors have described simultanagnosia as a reduction of the useful visual field (Bay, 1953; Thaiss & de Bleser, 1992; Tyler, 1968). More recently, Shalev, Humphreys, and Mevorach (2004) reported a restricted spatial area of attention in a simultanagnosic patient, which could be enlarged by a large preceding visual prime. Similarly, Michel and Henaff (2004) used crowding, counting, attentional tracking, and other tests to determine that a simultanagnosic patient had a "shrinkage" of the attentional visual field despite normal visual fields on perimetry.

To test the idea that simultanagnosia may be explained by a restriction of spatial attention, we previously restricted the visual window of healthy participants to only a small region of space at any one time (Dalrymple, Bischof, Cameron, Barton, & Kingstone, 2010). Thus, we explored the ability of a literal spatial constriction of *vision* to function as an analogue of the spatial constriction of a *spotlight of attention* in simultanagnosia. This was accomplished using a computer-generated gaze-contingent technique that allowed healthy individuals to see only within a small window surrounding their current fixation. Gaze-contingent displays have been used in the past with a variety

of tasks, such as reading (McConkie & Rayner, 1975), visual search (Pomplun, Reingold, & Shen, 2001), face perception (Caldara, Zhou, & Miellet, 2010), and scene exploration (Loschky, McConkie, Yang, & Miller, 2005). It is important to note that we do not suggest that a literal shrinkage of the visual field in healthy individuals is equivalent to a shrinkage of the attentional window in simultanagnosia. Rather, by restricting the visual window in healthy individuals, we are reducing the spatial area within which they can attend to a stimulus. Thus, those viewing scenes through a restricted window of vision and individuals with simultanagnosia are similarly restricted in terms of the spatial area from which they can gather visual information for the experimental task. Using this computer-generated gaze-contingent technique, we first found that healthy individuals viewing hierarchical letters through a restricted window of vision showed a similar difficulty in reporting global letters as simultanagnosic patients, and, as in simultanagnosia, their performance improved as the local elements became smaller or more densely packed (Dalrymple et al., 2010). This was the first indication that restricting the visual window of healthy individuals could be a useful tool for exploring at least some aspects of the simultanagnosic experience.

This possibility has led to a new set of questions: How do simultanagnosics view more complex and natural stimuli like social scenes, and how well does a restricted window of vision replicate their behaviour in these settings? We asked a simultanagnosic patient (S.L.) to describe social scenes while we monitored her eye movements (Dalrymple, Birmingham, Bischof, Barton, & Kingstone, 2011a). Unlike healthy subjects, who allocate a large proportion of fixations to the eyes of people in these scenes, S.L. distributed her gaze more evenly across the scenes, on objects, and on the heads and bodies of people. This contrasts with patients with parieto-occipital brain damage but no simultanagnosia, who show the normal tendency to fixate the eyes of people in scenes (Birmingham, Bischof, & Kingstone, 2007, 2008a, 2008b; Dalrymple et al., 2011a; Smilek, Birmingham, Cameron, Bischof, & Kingstone, 2006). Critically, reduced fixations on

the eyes that was apparent in S.L.'s behaviour was also found in our gaze-contingent simulation of simultanagnosia in healthy subjects: Individuals viewing social scenes through a restricted window of vision made abnormally few fixations on the eyes of people in the scenes.

While a restricted window of vision again appeared to capture the behavioural pattern of the restricted window of attention in simultanagnosia, it is unclear *why* these constraints should lead to reduced fixations on the eyes. One straightforward explanation is that it reflects a problem with the allocation of spatial attention in the restricted window condition and in simultanagnosia—a limitation that restricts the ability to locate and fixate the eyes of the people in the scenes. An alternative explanation is more in line with a suggestion by Birmingham et al. (2008b) who proposed that observers look at eyes of people in order to understand their attentional states. This account was tested by Birmingham et al. (2008b) by asking observers to either describe scenes that contained people or to infer the attention of people in scenes. The results indicated that observers made a significantly greater number of fixations on the eyes in the infer attention condition than when they were asked simply to describe the scenes. Thus an alternative explanation for why simultanagnosic patients and healthy individuals in the gaze-contingent paradigm reduce their fixations on the eyes is that the attentional states of the people in the scene are not prioritized when they are asked to describe a scene when much of the scene is not perceived. In other words, patients and healthy gaze-contingent participants *can* locate and fixate the eyes, but when scene perception is restricted, the eyes of people in the scenes are not prioritized because they provide little information about the content of the scene.

To distinguish between these two alternatives, in this study we monitored the eye movements of simultanagnosic patient G.B., healthy individuals viewing scenes through a gaze-contingent

window, and healthy individuals viewing scenes in an unrestricted viewing condition, while they performed two separate tasks (as in Birmingham et al., 2008b). In the first task they described social scenes, in part to replicate our previous findings that patients with simultanagnosia and healthy subjects with a restricted viewing window tend not to look at the eyes of people in social scenes. Replication here is important in and of itself because it confirms the unusual looking behaviour in a second patient with simultanagnosia and demonstrates the reliability of our findings with healthy subjects using the gaze-contingent display.[1] In the second task, which is novel to this study and included specifically to determine *why* simultanagnosics exhibit this unusual behaviour, the same individuals were asked to infer the attentional states of people in the scenes. Birmingham, Bischof, and Kingstone (2009) have demonstrated that the tendency to fixate eyes of people in scenes is not due to bottom-up saliency cues, but rather due to a goal-directed search for social attention information (see also Birmingham & Kingstone, 2009, for a full review on social attention and the importance of gaze information). Thus, this task was chosen because explicitly asking individuals to infer attentional states of people in social scenes increases the importance of the eyes to the normal healthy observer. If patients with simultanagnosia *cannot* find the eyes to use them to infer attention of people in scenes, they should make equally low proportions of fixations on the eyes of people in scenes in either task. However, if simultanagnosic patients are capable of locating and using the eyes to infer attention, but do not usually prioritize attentional states in scene viewing, then we should see an increase in fixations when the patients are asked explicitly to infer attentional states compared to when asked to describe the scenes. Finally, if the simultanagnosic behavioural results are related to an inability to process more than a small portion of the scenes at one time,

---

[1] The importance of replication has never been more important in the field of cognitive and social neuroscience, a point that has recently been driven home in the Special Issue of the *Perspectives on Psychological Science* (2012) Volume 7, Number 6.

there should be strong parallels with the behaviour of healthy subjects when they view scenes through a restricted viewing window.

The results will have important implications for our understanding of the nature of the attentional deficit in simultanagnosia. First, they will provide an index of the top-down control of selection of visual information in simultanagnosia. Traditionally, simultanagnosics have been described as having visual elements enter and exit awareness with little or no control over their visual experience (Holmes & Horrax, 1919; Rizzo & Hurtig, 1987). This claim is consistent with the possibility that simultanagnosics allocate few fixations on the eyes of people in scenes when asked to describe scenes because they are unable to locate the eyes. However, increased fixations on the eyes of people in scenes when explicitly inferring attentional states would indicate that simultanagnosics do in fact have top-down control over their selection of visual information, a finding that would contradict the view that their visual experience is mostly controlled by stimulus-driven information. It would also indicate that, unlike healthy individuals, social attention is not prioritized in simultanagnosia when merely describing a scene. Second, testing the validity of a spatially restricted window of vision as an analogue for simultanagnosia across different tasks will speak to the debate between object-based versus space-based theories of simultanagnosia. Similarities between simultanagnosic scanning behaviour and that of healthy individuals viewing scenes through a spatially restricted window of vision would support a space-based theory of simultanagnosia. Critically, it would also suggest that a restricted visual field for attentional selection may contribute to a decline in prioritization of social attention in simultanagnosia when other task demands compete for limited attentional resources.

Method

Participants

Patient G.B.. G.B. is a 32-year-old man with posterior reversible leukoencephalopathy (PRES).

Magnetic resonance imaging (MRI) revealed localized abnormal hyperintense fluid-attenuated inversion recovery (FLAIR) signals in medial occipital lobes and parietal lobes, mainly involving subcortical U fibres (see Figure 1). Goldmann perimetry initially revealed paracentral right inferior homonymous relative scotoma involving the central 5 degrees. He had a dressing apraxia and some problems with visual search tasks. Reading was normal in fluency. He did well on a subset of items from the famous faces test ruling out the possibility of prosopagnosia. His performance on a selection of items from the Hooper Visual Organization Test revealed a few errors suggesting inattention to detail. He could copy geometric designs but showed poor planning in clock drawing, with numbers placed inside and outside the ring. Previous neuropsychological testing on the Wechsler Adult Intelligence Scale–Third Edition (WAIS–III) suggested average to superior premorbid intelligence, based on verbal comprehension, vocabulary, and general knowledge. G.B. did well on numerical calculations and repetition, but had problems with left–right orientation and visual memory. Scores from his neuropsychological evaluation can be found in the supplementary online materials.

Standard diagnostic tools for simultanagnosia are the Boston Cookie Theft picture (Goodglass & Kaplan, 1983), which patients tend to describe in a piecemeal manner (Rizzo & Vecera, 2002), and the overlapping figures test, which consists of overlapping line drawings of familiar objects. Simultanagnosics tend to report seeing only one of the overlapping figures at one time (Rafal, 2003). Consistent with previous reports of simultanagnosia, G.B. described the Boston Cookie Theft picture in a piecemeal manner, slowly, and laboriously, despite having seen the picture before. For the overlapping figures test, G.B. had difficulty identifying more than one element at a time. For example, he said he saw a basket and that he thought that there were items in the basket, but that he could not identify the items. Eventually he identified a clock (which was actually a wrist watch) and much later a key, but
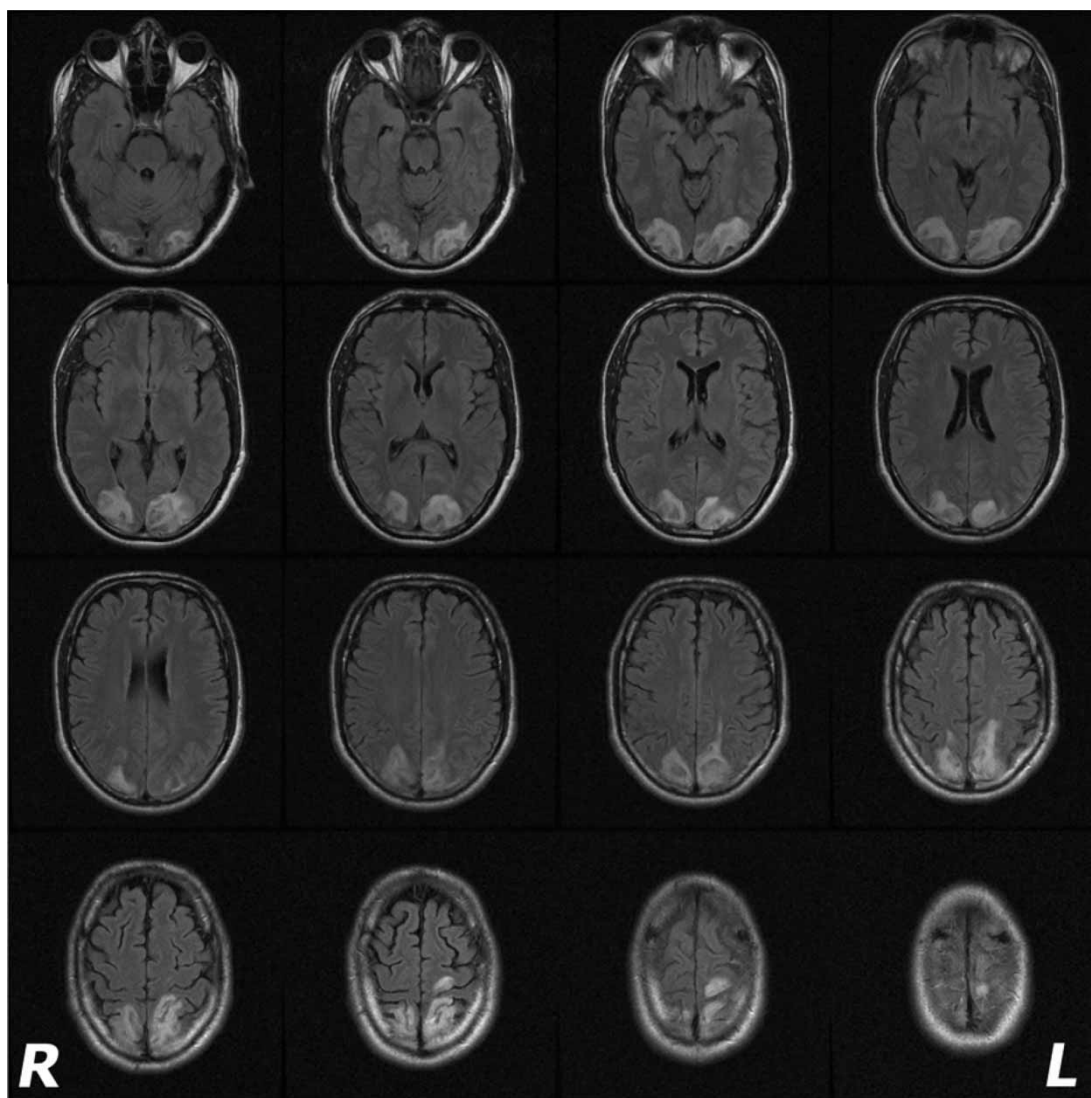
Figure 1. *Magnetic resonance imaging (MRI) scans of patient G.B.*

much of this was done through inferences about shapes that he could discriminate. He also identified a cup, but eventually decided that it may in fact be something else (it was actually a pipe). He never identified the light bulb.

Recent empirical work has revealed consistencies across simultanagnosics on other tasks. For example, when viewing hierarchical letters, which are global letters made up of several repetitions of a local letter, patients with simultanagnosia tend to report seeing the local letters, but not the global letters (Clavagnier et al., 2006; Karnath et al., 2000; Shalev et al., 2004). Performance on global letters tends to improve when interelement spacing is reduced (Dalrymple et al., 2009; Dalrymple et al., 2007). We tested G.B. with hierarchical letters of different sizes and densities. He performed poorly at identifying the global form when the letters were constructed of sparse local elements, but was near perfect when they were made up of local

Table 1. *G.B.'s accuracy for naming global and local levels of hierarchical letters*

Name	Size	Density Sparse	Medium	Dense
Global	Small	45	64	91
	Medium	64	91	82
	Large	82	91	82
Local	Small	100	100	100
	Medium	100	100	100
	Large	100	100	100

Note: Accuracy in %.

elements that were densely packed. He performed perfectly at identifying the local letters for all size/density configurations (Table 1).

G.B. also performed like a previously reported simultanagnosic (S.L.) on a hierarchical faces test. Hierarchical faces are faces made up of a series of objects (e.g., a face made up of fruit with a pear for a nose and berries for eyes). When viewing these faces, S.L. reported seeing just the face, or the objects that make up the face, but not both (Dalrymple et al., 2007). Similarly, G.B. identified either the elements that made up the face (e.g., berries), or the face, but never identified both in the same image.

Despite his simultanagnosia, G.B. performed normally when asked to name real objects, body parts, and line drawings of inanimate and animate objects. He had average abilities for discrimination of visual detail, pattern recognition, and picture arrangement. He did well on degraded stimuli tasks and on object vision generally. At the time of testing, his visual field defects were much reduced (Figure 2). He still had visuospatial orientation difficulties and dressing apraxia.

Comparison groups. Our primary research questions are specific and concern (a) the change, if any, in G.B.'s fixations on the eyes during the describe task versus the infer attention task, and (b) whether G.B.'s eye movement behaviour during these tasks can be mimicked by restricting the spatial extent of the visual window of healthy individuals performing the same tasks. To anticipate our results, the answer is that healthy individuals will appear "simultanagnosic" on the present task when the spatial extent of the visual window is constrained. Thus, we included two comparison groups in this study. The "full-view group" consisted of healthy individuals who viewed the scenes naturally and allowed for a measure of normal fixation allocation in each task. The "gaze-contingent group" consisted of healthy
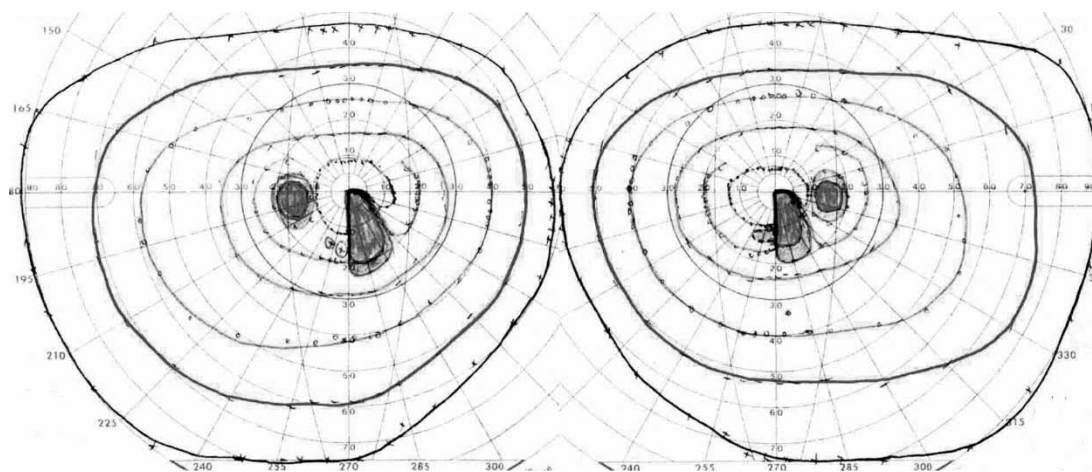

Figure 2. *Goldmann perimetry for G.B., showing small homonymous paracentral scotoma in the right lower quadrant.*

individuals who viewed the scenes through a computer-generated aperture and allowed for the evaluation of this spatial restriction of vision as an analogue for simultanagnosia across tasks. It is worth noting that in a previous investigation of eye movement behaviour in simultanagnosia (Dalrymple et al., 2011a), we tested brain-damaged control participants and found that it is simultanagnosia, and not parietal damage per se, that leads to reduced fixations on the eyes when describing scenes.

Full-view group. Participants (*n* = 12, 5 male) were undergraduate students at the University of British Columbia who ranged in age from 19 to 28 years (mean = 23 years). All participants in this and the subsequent group reported normal or corrected-to-normal vision and gave informed consent prior to participation in the experiments, which were performed in accordance with the ethical guidelines of the University of British Columbia.

Gaze-contingent group. Participants (*n* = 14, 6 male) were undergraduate students at the University of British Columbia who ranged in age from 19 to 56 years (mean = 25.6 years).

Stimuli and apparatus

Full colour images were taken with a digital camera in different rooms in the Psychology building at the University of British Columbia. Image size was 36.5 × 27.5 (cm) corresponding to 40.1° × 30.8° at the viewing distance of 50 cm, and image resolution was 800 × 600 pixels. Sixteen scenes were used in the present experiment. Each scene contained three persons. All scenes were comparable in terms of their basic layout: each room had a table, chairs, objects, and background items (e.g., see Figure 3a).

Eye movements were monitored using the desk-mounted EyeLink 1000 eye tracking system (SR Research Ltd., www.eyelinkinfo. com). The Eyelink 1000 has a temporal resolution of 1 ms and a spatial resolution of 0.5°. It records data indicating the location of gaze in pixel coordinates. Before any analysis was
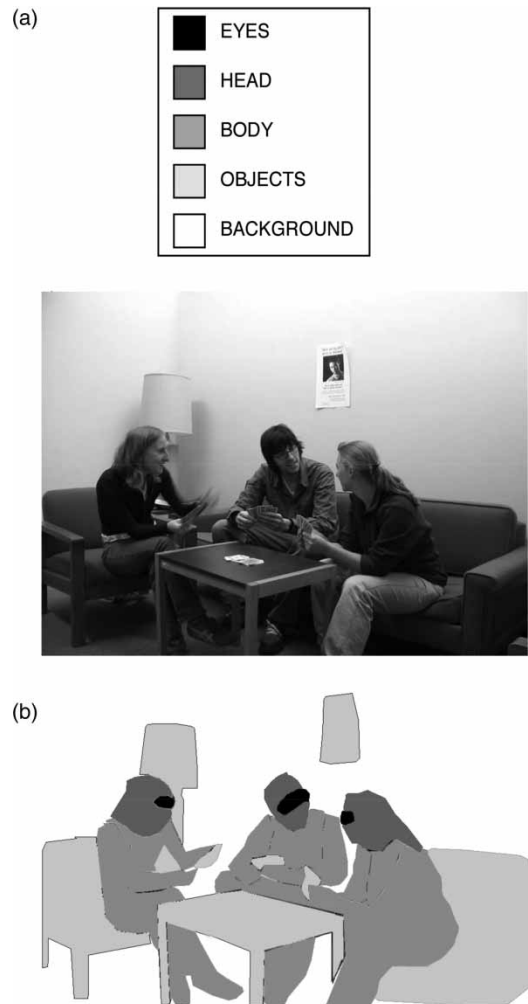

Figure 3. *Example of scene stimuli (a) and regions of interest (b).*

carried out, these data were parsed into fixation and saccade events (and blinks) using the EyeLink software. The event parser identifies epochs in the data file where a saccade is occurring by calculating the distance between gaze position in different samples and implementing motion, velocity, and acceleration thresholds. The online saccade detector of the eye tracker was set to detect saccades with an amplitude of at least 0.15°, using an acceleration threshold of 8,000° s^2 and a velocity threshold of 30° s. A fixation is defined as any event that was not a

saccade or a blink. Two computers were used in the experimental set-up and were connected to each other via Ethernet, allowing for real-time transfer of saccade and gaze position data. The experimenter's computer collected the data from the eye tracker and displayed an image of the participant's eye and calibration information. The display computer presented the stimuli and recorded keypresses.

Subjects in the gaze-contingent group viewed the scenes through a $2° \times 2°$ square aperture centred on their current fixation point. This size was chosen based on previous studies modelling simultanagnosic behaviour with a gaze-contingent aperture (Dalrymple et al., 2011a; Dalrymple, Birmingham, Bischof, Barton, & Kingstone, 2011b; Dalrymple et al., 2010). Outside of the portion of the scene visible through the aperture, the screen was white. As the subject's fixation moved, the aperture moved with the fixation point. Subjects in this group underwent a short training session to familiarize them with the gaze-contingent window: They were instructed to start from a circle at the centre of the screen labelled "Start" and to follow a line with the gaze-contingent window from that circle until they reached a second circle labelled "End". They were then instructed to freely search the screen for a hidden object on the screen. This task was irrelevant to the experimental task and was designed to teach them how to control the gaze-contingent aperture. Once they located the hidden object and felt comfortable with the apparatus, the experiment began. Eye-monitoring studies can be demanding for participants, in particular those in the gaze-contingent condition. Six participants were excluded because they failed to complete the study (four from the gaze-contingent condition), and two because their fixation proportions fell more than 2 standard deviations from the mean.

Procedure
Subjects were seated 50 cm from the screen of the display computer with their chin supported by a rest. Eye movements were recorded from the left eye. The eye monitor was calibrated and validated using a 9-dot array. Subjects were then asked to fixate a dot at centre-screen while the experimenter corrected for drift in gaze position. Once the dot was fixated, the experimenter initiated the onset of the scene image by a keypress.

There were two tasks (describe and infer) and two sets of eight scenes (A and B). Scenes were randomly assigned to Set A or B. In the describe task, subjects were asked to verbally describe each scene (i.e., "Please describe the scenes out loud"). In the infer task, subjects were asked to verbally describe where people in the picture were directing their attention (i.e., "Please describe out loud where the people in the scenes are directing their attention"). Trials were blocked by task, and each task was performed with both scene sets, for a total of four blocks (and a total of 32 trials). Subjects alternated between tasks (e.g., describe Set A, infer Set B, describe Set B, infer Set A), with the assignment of the first task counterbalanced across subjects. G.B. first performed the describe condition. Although the order of scene sets was fixed (i.e., ABBA or BAAB), scenes were presented in random order within a block. All verbal responses were given concurrently with scene viewing and were recorded using a digital voice recorder. Subjects terminated their trials by keypress, initiating the next trial. They had a maximum of 3 minutes to view each scene, but rarely reached this time limit.

Analysis
For each image, an outline was drawn around each region of interest (e.g., "eyes"), and each region's pixel coordinates and area were recorded. We defined the following regions: eyes, head (excluding eyes), body (including arms, torso, and legs), foreground objects (e.g., tables, chairs, objects on the table) and background (e.g., walls, shelves, items on the walls). Figure 3b illustrates these regions for one scene. To compensate for the different sizes of these regions, we computed area-normalized fixation proportions (Birmingham et al., 2008b; Smilek et al., 2006), by first dividing the number of fixations in each region by the area of the region, separately for each image and each participant, and then computing proportions based on these normalized data.

Given our specific research question and *a priori* hypotheses, we focused our analysis on the eyes region only. We used two-tailed paired *t* tests to compare proportions of fixations on the eyes in the describe task to those in the infer task separately for each group (full-view, gaze-contingent, and G.B.). We then used Crawford, Garthwaite, and Howell (2009) modified *t* tests using SINGLIMS software (Crawford et al., 2009; Sokal & Rohlf, 1995) to compare G.B. to the full-view and gaze-contingent groups, respectively, in each task (describe and infer). All *p* values were compared to ≤.05.

Results

Full-view, gaze-contingent, and G.B.
Figure 4 shows the mean proportions of fixations on the eyes region in the describe and infer conditions for each group. Two-tailed paired *t* tests revealed that all groups made a significantly larger proportion of fixations on the eyes in the infer condition than in the describe condition: full-view, $t(7) = -4.37$, $p = .003$; gaze-contingent, $t(8) = -3.57$, $p = .007$; G.B., $t(15) = -5.44$, $p < .001$. Figure 5 shows representative scan patterns for each group and task.
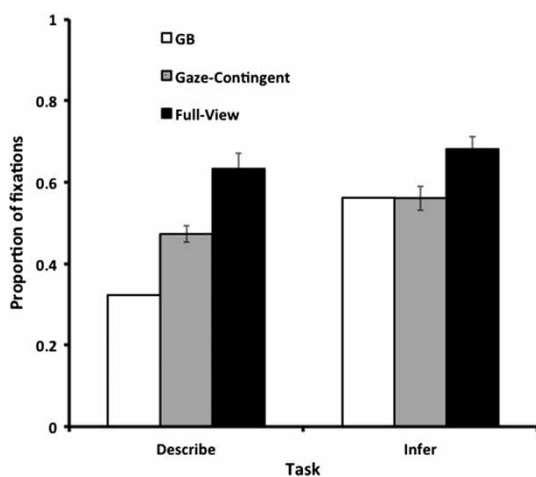
Figure 4. *Fixation proportions on the eyes for each group for the describe and infer tasks. Fixation proportions are normalized to the size of each region. Error bars represent standard error.*

Full-view versus gaze-contingent
We used a two-way analysis of variance (ANOVA) with factors of group (full-view vs. gaze-contingent) and task (infer vs. describe) to compare the full-view group to the gaze-contingent group in terms of their fixations on the eyes. There was a main effect of group, $F(1, 35) = 11.83$, $MSE = .173$, $p = .003$, and a main effect of task, $F(1, 35) = 24.23$, $MSE = .042$, $p < .001$, but no Group × Task interaction, $F(1, 35) = 1.97$, $MSE = .003$, $p = .180$. The effect of group was due to a larger proportion of fixations on the eyes in the full-view than in the gaze-contingent group. The effect of task was due to both groups looking at the eyes more when inferring the attention of people in the scenes than when asked to describe the scenes. The lack of interaction indicates that fixations on the eyes for both groups varied by task in the same way.

G.B. versus full-view group
We used Crawford et al.'s (2009) modified *t* tests (Crawford et al., 2009; Sokal & Rohlf, 1995) to compare G.B. to the full-view group. These tests revealed that G.B. had a significantly smaller proportion of fixations on the eyes than the full-view group in the describe condition, $t(7) = -2.62$, $p = .034$, but that he did not differ from them in the infer condition, $t(7) = -1.17$, $p = .280$. That is, when the task involved inferring the direction of attention of the people in the scenes, G.B. was just as likely to look at the eyes as were the full-view group.

G.B. versus gaze-contingent group
Crawford et al.'s (2009) modified *t* test revealed that G.B. did not differ significantly from the gaze-contingent group in either the describe, $t(8) = -2.18$, $p = .061$, or the infer conditions, $t(8) = 0.02$, $p = .984$.

Verbal reports
Verbal reports from G.B. and participants from the gaze-contingent and full-view groups (Table 2) were transcribed and inspected to ensure that participants were capable of performing the tasks and that they understood the
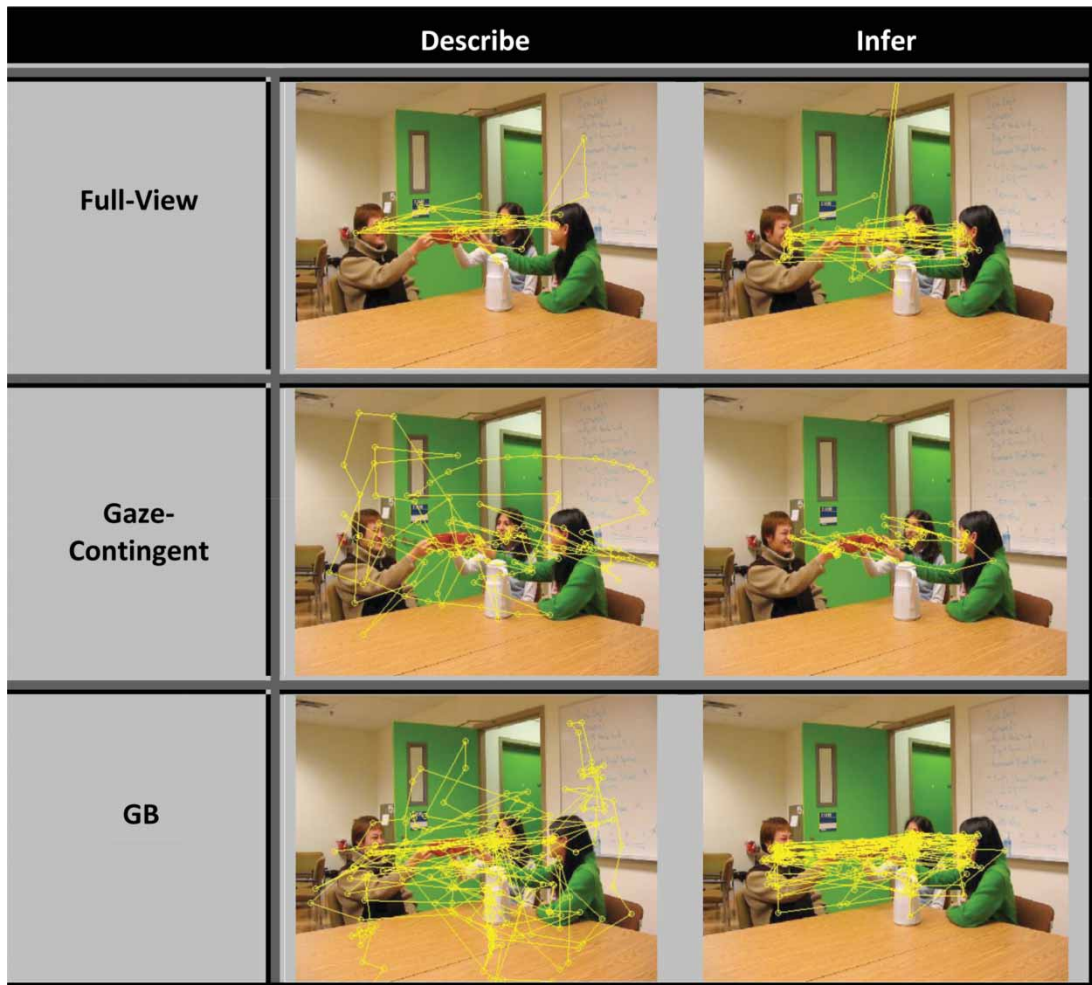
Figure 5. *Representative scan paths for the full-view and gaze-contingent groups and for G.B. for both the describe and infer tasks. Circles represent fixations. Lines represent movements from one fixation to the next. To view a colour version of this figure, please see the online issue of the Journal.*

difference between them. These reports clearly indicate that when asked to describe a scene, participants were able to synthesize meaning from the elements that they perceived. They also indicate that participants understood the tasks: Their descriptions during the "infer attention" task show an understanding that attention can be inferred by looking at the eyes of the individuals in scene. These reports also corroborate the eye movement measures: Participants look at and discuss the eyes more during the "infer attention" task than during the "describe the scenes" task.

Trial durations

On average, the full-view group spent 25.96 s (SD = 11.81) describing scenes and 28.51 s (SD = 14.01) inferring the attentional states of people in the scenes. The gaze-contingent group spent on average 53.37 s (SD = 24.54) describing the scenes and 49.26 s (SD = 10.62) inferring the attentional states of people in the scenes. G.B. spent on average 60.21 s describing the scenes and 78.68 s inferring the attentional states of people in the scenes. The gaze-contingent group had significantly longer trial durations overall than

the full-view group, $F(1, 35) = 13.00$, $p = .002$, but there was no significant effect of task, $F(1, 35) = 0.04$, $p = .836$, or Task × Group interaction, $F(1, 35) = 0.80$, $p = .383$, indicating that both groups had equal trial durations across tasks. Similar to the gaze-contingent group, G.B.'s trial durations were significantly longer than those of the full-view group for both tasks: describe, $t(8) = 2.75$, $p = .025$; infer, $t(8) = 3.40$, $p = .009$. However, while G.B. did not differ from the gaze-contingent group in terms of time spent describing scenes, $t(8) = 0.26$, $p = .798$, he did spend significantly longer inferring the attentional states of people in the scenes than this group, $t(8) = 2.63$, $p = .030$. Verbal reports (Table 2) confirm that all participants are capable of performing the tasks, so increased trial durations for G.B. and the gaze-contingent participants relative to the full-view group appear to reflect the increased time needed to gather information when restricted to seeing only a small portion of the scene at one time.

Discussion

Here we highlight three main aspects of our results. First, we replicate a previous finding from another simultanagnosic patient S.L. (e.g., from Dalrymple et al., 2011a, 2011b): When describing social scenes, patient G.B. makes few fixations on the eyes of people in the scenes. This is important because it supports the suggestion that this unusual behaviour is related to simultanagnosia, rather than being an anomalous finding from a single case. Second, when engaged in a task that explicitly requires G.B. to infer attention, a task that encourages eye fixations in healthy subjects, he does fixate them, and with a proportion of fixations that is no different from that of healthy controls performing the same task. Thus, like healthy subjects, G.B. shows an understanding that the eyes are an important source of information for attentional states and uses this understanding to allocate his fixations in a goal-directed way. G.B.'s understanding of the importance of the eyes for inferring attentional states is evident in his verbal reports (Table 2). His behaviour indicates that he can locate and use the eyes to infer attention, but when simply describing scenes he does not give others' attentional states priority to the degree that healthy individuals do. The notion has been raised that, when asked to describe social scenes, G.B. may be using parts of the scenes other than the eyes to gather information about the attentional states of people in the scenes. However, G.B.'s tendency to increase fixations on the eye regions when explicitly asked to infer attentional states suggests that, like healthy individuals, he does in fact gather information about attentional states by looking at eyes. Thus, his reduced fixations on the eyes when describing scenes suggests that he is prioritizing information other than attentional states during this task. These findings also indicate that G.B. can execute top-down control over his selection of visual information.

The third main aspect of our results is that G.B.'s behaviour is largely reproduced by healthy subjects viewing the scenes through a restricted window of vision. Compared to subjects viewing scenes without any restriction, these subjects, like G.B., showed fewer fixations on the eyes when describing the scenes. However, as with subjects viewing without restriction, subjects in the gaze-contingent window condition showed an increase in fixations on eyes when inferring the direction of attention of people in scenes. This discovery lends new and converging support to a restricted window of vision as a useful model for exploring simultanagnosic behaviour with healthy individuals.[2] Importantly, this speaks to the nature of the attentional deficit in simultanagnosia, supporting

---

[2] The present finding is especially remarkable because the gaze-contingent group consisted of young controls (mean age 25.6 years), demonstrating that visual restriction, and not age, is the critical factor. We have previously shown that the fixation behaviour of healthy individuals who were matched in age to the simultanagnosic patient S.L. (49 years) did not differ from the fixation behaviour of a group of young controls (mean age 22 years), further suggesting that age is not critical to our results (Dalrymple et al., 2011a).

Table 2. *Verbal reports from G.B. and randomly selected participants from the full-view and gaze-contingent groups*

Group	Describe	Infer
Full-view	Three students, probably in the [psychology building], are drinking tea or something, and they're doing cheers. They look happy.	Ummm . . . this picture looks like the young man looking . . . looking at the woman in the green shirt and the lady in the green shirt . . . hard to say . . . looks like she's looking . . . maybe at the—at the young man . . . and then the la- the girl in the middle looks like she's looking at the cups in general.
Gaze-contingent	K, there's . . . uh (long pause) two women, one's wearing a green shirt, it looks like . . . um . . . smiling . . . (long pause) . . . holding a cup . . . uh . . . in a room . . . with a green door, looks like (long pause). Um . . . (long pause). There's a man as well, I think they're cheering or something, putting their cups together. Uh . . . (long pause). There's a lot of green. So it looks like there's three people in the room—that's all I see, anyway.	So I see three cu-cups coming together so they look like they're . . . banging their cups together so that's what their direction is atten- towards, all three of them it seems. K . . . yeah . . . yeah, they're all looking towards the same spot.
G.B.	(Long pause) There's three people doing cheers, with um . . . coffee cups and there's um a coffee (pause) carafe on the table . . . and umm . . . (pause) three people, there's a woman in a red- a green shirt (pause) and another woman (pause) and um . . . (pause) a man and (pause) a green door (pause) to the room that's propped open (pause) and there's stuff written on a whiteboard, behind um . . . (pause)(inaudible) . . . I can't see what it is . . . (long pause).	(Long pause) OK so I, I . . . I see who's directing their attention at (pause) at . . . what or whom and I see (inaudible) in fact (pause) um . . . (pause) it looks to me like all three people are, are directing their attention at the . . . at their cups, but, we- they're doing cheers and they're all looking at their . . . at their cups (pause) and um . . . (pause) that's because I'm looking at their eyes (pause) and they don't look like they are angled in a way so they could be looking at making um eye contact with . . . any of the other people . . . (pause) and I also think it's just normal when you do cheers you . . . watch the cups so you don't spill anything or . . . knock it too hard.

Note: Reports were recorded during the trials depicted in Figure 5.

space-based, as opposed to object-based, account of the disorder.

The data from our full-view control subjects fit nicely with previous findings by Birmingham et al. (2008a) who found that task modulated the fixation behaviour of healthy subjects in a similar way to what we found in the present study. This replication of the Birmingham et al. results is particularly interesting because it occurred despite some key methodological differences between our study and theirs. Birmingham et al. allowed subjects to view the scenes for only 15 seconds, at which point they were asked to describe what they saw or infer the attention of people in the scenes (i.e., when the scene was no longer visible). In contrast, subjects in our study viewed scenes for up to 3 minutes and were asked to describe the scenes or infer attention while they viewed the scene. This means that subjects allocate their fixations in a similar way for both relatively short and prolonged scene viewings, and whether they provided immediate or delayed report. This attests to the robustness of the tendency for subjects with normal vision to look at eyes in social scenes and supports the idea that the findings in this paradigm are capturing a behavioural profile that generalizes across situational complexities.

Collectively, our findings suggest that although patient G.B. is capable of looking at the eyes when necessary to the task (i.e., during the infer task), he tends not to look at the eyes when it is not explicitly required of him (i.e., during the describe task). This supports the hypothesis that the attentional states of others are not a high priority for simultanagnosics when they are describing a scene. The present data, combined with previous and current findings from modelling simultanagnosia in healthy individuals, can offer insight into why these patients may reprioritize the importance of the eyes. We previously suggested that some aspects of simultanagnosia may be modelled as a restriction of the spatial window of attention, a type of "attentional tunnel vision". This notion dovetails with initial reports documenting a reduction of the useful visual field in simultanagnosia (Bay, 1953; Thaiss & de Bleser, 1992; Tyler, 1968), as well as more recent findings suggesting a

"shrinkage" of the attentional visual field in simultanagnosia despite normal visual fields (Michel & Henaff, 2004), and the enlargement of an abnormally small spatial area of attention in simultanagnosia through the use of large preceding visual primes (Shalev et al., 2004). Our prior work has shown that healthy subjects viewing scenes through a restricted window of *vision* behave much like simultanagnosic patients in processing hierarchical letters (Dalrymple et al., 2010) and when describing social scenes (Dalrymple et al., 2011a, 2011b). Here we demonstrate that the parallels between limiting the spatial extent of the visual window in healthy subjects and the attentional restriction in simultanagnosia extend to the task of inferring the attention of people in social scenes.

In line with results from a previous study (Dalrymple et al., 2011a), the describe task data from G.B. and from our gaze-contingent group suggest that when only one element of a scene can be processed at a time, the most efficient way to describe the scene is to scan more elements of the scene rather than revisiting a single region such as the eyes. This contrasts with healthy subjects with unrestricted vision, who can allocate the majority of their fixations to the eyes of people and still describe the rest of the scene using gist, or vision at a glance from information in their parafoveal vision (Hochstein & Ahissar, 2002). However, when the task shifts to inferring attention, all groups realize that the primary source of information lies in the eyes, which can still be located and fixated despite restrictions in the span of either vision or attention. Critically, this finding indicates that, like the non-brain-damaged participants who viewed scenes through a restricted window of vision, G.B. maintains some top-down control over his fixations.

This conclusion is consistent with previous work that has shown top-down processing strategies in another individual (patient J.J.) with simultanagnosia (Jackson, Swainson, Mort, Husain, & Jackson, 2009). For example, in a pop-out search task, where a target with unique features is presented among distractors that do not contain those features, J.J. used a systematic,

item-by-item search strategy to find the target. Our findings, and those of Jackson et al. (2009), provide convergent evidence against previous suggestions that the visual experience of simultanagnosics is largely controlled by bottom-up factors such as object saliency (Holmes & Horrax, 1919; Rizzo & Hurtig, 1987).

Data from the gaze-contingent group show that restricting the useful visual field of healthy individuals leads to similar fixation patterns to those that are recorded with simultanagnosic patients who can only see one thing at a time, whether describing social scenes or inferring the direction of attention of people in social scenes. However, it is important to be clear about the limitations of the restricted window of vision as a model of simultanagnosia. We do not claim that a restricted window of vision models all aspect of the simultanagnosic deficit. Rather, we are using the restricted window of vision to reduce the spatial area to which healthy subjects can attend, akin to the reduction of the visual window within which attentional selection can operate in simultanagnosia. In fact, one strength of this method is that this simple manipulation can lead to complex scanning behaviours in healthy adults that are similar to those seen in simultanagnosia across a variety of tasks (e.g., global/local processing, Dalrymple, et al., 2010; recovery from simultanagnosia, Dalrymple et al., 2011b; describing scenes, Dalrymple et al., 2011a; and inferring attention of people in scenes as reported here).

It is important to note that although G.B. and the gaze-contingent group fixate the eyes to a similar degree when describing the scenes, there was a trend for these groups to differ from each other in this task, with the gaze-contingent group fixating the eyes slightly more than G.B. However, we have previously shown that the degree of fit between simultanagnosic behaviour and the gaze-contingent model is influenced by the size of the window used in the gaze-contingent paradigm (Dalrymple et al., 2011b)—a smaller window tends to lead to fewer fixations on the eyes and might have been a better fit for patient G.B.'s behaviour. Ultimately, it is not the exact proportion of

fixations that is of critical importance here. Rather, it is the tendency for both G.B. and the gaze-contingent group to show abnormally low fixation proportions on the eyes when describing scenes and to increase that proportion of fixations when inferring attention that makes this a useful model of social scene viewing in simultanagnosia.

One question that may arise from this and previous studies of complex scene viewing in simultanagnosia is to what degree simultanagnosics are able to describe and understand the scenes that they are viewing. Verbal descriptions from G.B. (Table 2) and from previous scene-viewing experiments with simultanagnosia patient S.L. (Dalrymple et al., 2011a, 2011b) indicate that patients with simultanagnosia are capable of piecing together elements of a scene to describe the scene as a whole. This process is often slow and laborious, and not always perfectly accurate, but may result in a surprising degree of coherence. Rizzo and Vecera (2002) articulate how it is possible for an individual with simultanagnosia to report multiple elements of a scene in a seemingly coherent fashion: " . . . object recognition can proceed economically from just a few key features, and identifying a person or object is not the same as seeing him, her, or it all at once" (p. 164). Patients with simultanagnosia suffer from a visual disorder that may (and often does) leave their cognitive faculties intact (e.g., see G.B.'s case report). Thus patients are able to compensate for their visual disorder by consciously synthesizing meaning from what they see.

We would also like to acknowledge that an anonymous reviewer flagged for our attention the fact that these data give rise to a separate but interesting question for future investigation concerning whether G.B.'s top-down attention to the eyes when inferring attentional states interferes with his ability to derive meaning from the scene. Jackson et al. (2009) suggest that goal-directed behaviour in simultanagnosia comes at the expense of competing stimuli that fall outside of the immediate task demands. Applied here, this suggests that when explicitly asked to determine the attentional states of people in the scenes, G.B. may have

sacrificed some understanding of the scene as a whole in order to perform the task. However, it is also possible that a comparable loss may have occurred for G.B. during the "describe the scenes" task, and that this loss may not be specific to simultanagnosia: It is well known in attention research that selection for some items comes at a cost of reduced processing for others (Broadbent & Broadbent, 1987). Indeed, the content of G.B.'s verbal reports from the infer attention and describe the scenes tasks were convergent with those provided by the healthy participants in both the full-view and the gaze-contingent conditions. Future investigations using different task instructions could be used to determine to what extent goal-oriented behaviour impedes additional information processing in simultanagnosia.

Importantly for our primary research question, our results indicate that patients strategize their scene exploration based on task demands. When asked to describe a scene, patients use goal-directed behaviour to distribute their fixations across the scene to gather information about the scene. When asked to infer the attentional states of people in the scenes, patients, like healthy controls, increase fixations on the eyes. These strategies were also employed by healthy individuals who viewed the scenes though a restricted window of vision, further indicating that it is likely to be the restricted window of visual information, not brain damage per se, that is driving this behaviour.

Seeing one object at a time, as in simultanagnosia and in gaze-contingent viewing, drastically affects how one visually explores the world. What our results suggest is that what could previously have been interpreted as a lack of control over exploratory eye movements (leading to an avoidance of the eyes of people in scenes when describing the scenes) can now be understood as a top-down effort to derive information from a scene. Ultimately we can conclude that simultanagnosic eye movements are task dependent. How they vary with the task is informative about the overall visual experience of patients with this profound disorder and about the direction of attention in situations with limited visual information.

Supplementary material

Supplementary Material is available via the "Supplementary" tab on the article's online page (http://dx.doi.org/10.1080/02643294.2013.778234).

REFERENCES

Bálint, R. (1909). Seelenlähmung des "Schauens", optische Ataxia, räumliche Störung der Aufmerksamkeit [Psychic paralysis of gaze, optic ataxia, and spatial disorder of attention]. *Monatsschrift für Psychiatrie und Neurologie, 25,* 51–81.

Bay, E. (1953). Disturbances of visual perception and their examination. *Brain, 76,* 515–550.

Baylis, G., Driver, J., Baylis, L., & Rafal, R. (1994). Reading of letters and words in a patient with Balint's syndrome. *Neuropsychologia, 32,* 1273–1286.

Birmingham, E., Bischof, W., & Kingstone, A. (2007). Why do we look at people's eyes? *Journal of Eye Movement Research, 1,* 1–6.

Birmingham, E., Bischof, W., & Kingstone, A. (2008a). Gaze selection in complex social scenes. *Visual Cognition, 16,* 341–355.

Birmingham, E., Bischof, W., & Kingstone, A. (2008b). Social attention and real world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61,* 986–998.

Birmingham, E., Bischof, W., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49,* 2992–3000.

Birmingham, E., & Kingstone, A. (2009). Human social attention: A new look at past, present, and future investigations. *Annual New York Academy of Sciences, 1156,* 118–140.

Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42*(2), 105–113.

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the "spotlight" reveals universal information use for face recognition. *PLoS One, 5*(3), e9708.

Clavagnier, S., Fruhmann Berger, M., Klockgether, T., Moskau, S., & Karnath, H.-O. (2006). Restricted

ocular exploration does not seem to explain simultanagnosia. *Neuropsychologia, 44,* 2330–2336.

Crawford, J., Garthwaite, P., & Howell, D. (2009). On comparing a single case with a control sample: An alternative perspective. *Neuropsychologia, 47,* 2690–2695.

Dalrymple, K. A., Birmingham, E., Bischof, W., Barton, J., & Kingstone, A. (2011a). Experiencing simultanagnosia through windowed viewing of social scenes. *Brain Research, 1367,* 265–277.

Dalrymple, K. A., Birmingham, E., Bischof, W., Barton, J., & Kingstone, A. (2011b). Opening a window on attention: Documenting and simulating recovery from simultanagnosia. *Cortex, 47,* 787–799.

Dalrymple, K. A., Bischof, W., Cameron, D., Barton, J., & Kingstone, A. (2009). Global perception in simultanagnosia is not as simple as a game of connect-the-dots. *Vision Research, 49,* 1901–1908.

Dalrymple, K. A., Bischof, W., Cameron, D., Barton, J., & Kingstone, A. (2010). Simulating simultanagnosia: Spatially constricted vision mimics local capture and the global processing deficit. *Experimental Brain Research, 202,* 445–455.

Dalrymple, K. A., Kingstone, A., & Barton, J. (2007). Seeing trees OR seeing forests in simultanagnosia: Attentional capture can be local or global. *Neuropsychologia, 45,* 871–875.

Damasio, A. (1985). Disorders of complex visual processing: Agnosias, achromatopsia, Balint's syndrome, and related difficulties of orientation and construction. In M.-M. Mesulam (Ed.), *Principles of behavioral neurology* (pp. 259–288). Philadelphia, PA: F.A. Davis Company.

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders*. Philadephia, PA: Lea and Febiger.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791–804.

Holmes, G., & Horrax, G. (1919). Disturbances of spatial orientation and visual attention, with loss of stereoscopic vision. *Archives of Neurology and Psychiatry, 1,* 385–407.

Jackson, G., Swainson, R., Mort, D., Husain, M., & Jackson, S. (2009). Attention, competition, and the parietal lobes: Insights from Balint syndrome. *Psychological Research, 73*(2), 263–270.

Karnath, H.-O., Ferber, S., Rorden, C., & Driver, J. (2000). The fate of global information in dorsal simultanagnosia. *Neurocase, 6,* 295–306.

Loschky, L., McConkie, G., Yang, J., & Miller, M. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition, 12*(6), 1057–1092.

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

Michel, F., & Henaff, M.-A. (2004). Seeing without the occipito-parietal cortex: Simultanagnosia as a shrinkage of the attentional visual field. *Behavioural Neurology, 15,* 3–13.

Moreaud, O. (2003). Balint syndrome. *Archives of Neurology, 60,* 1329–1331.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Pomplun, M., Reingold, E., & Shen, J. (2001). Investigating the visual span in comparative search: The effects of task difficulty and divided attention. *Cognition, 81,* B57–B67.

Rafal, R. (1997). Balint syndrome. In T. Feinberg & M. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 337–356). New York, NY: McGraw-Hill.

Rafal, R. (2003). Balint's syndrome: A disorder of visual cognition. In M. D'Esposito (Ed.), *Neurological foundations of cognitive neuroscience* (pp. 27–40). Cambridge, MA: MIT Press.

Riddoch, M., Chechlacz, M., Mevorach, C., Mavritsaki, E., Allen, H., & Humphreys, G. (2010). The neural mechanisms of visual selection: The view from neuropsychology. *Annals of the New York Academy of Sciences, 1191,* 156–181.

Rizzo, M., & Hurtig, R. (1987). Looking but not seeing: Attention, perception, and eye movements in simultanagnosia. *Neurology, 37*(10), 1642–1648.

Rizzo, M., & Vecera, S. (2002). Psychoanatomical substrates of Balint's syndrome. *Journal of Neurology, Neurosurgery & Psychiatry, 72,* 162–178.

Shalev, L., Humphreys, G., & Mevorach, C. (2004). Global processing of compound letters in a patient with Balint's syndrome. *Cognitive Neuropsychology, 22,* 737–751.

Smilek, D., Birmingham, E., Cameron, D., Bischof, W., & Kingstone, A. (2006). Cognitive ethology and exploring attention in real-world scenes. *Brain Research, 1080,* 101–119.

Sokal, R., & Rohlf, J. (1995). *Biometry*. San Francisco, CA: W.H. Freeman.

Thaiss, L., & de Bleser, R. (1992). Visual agnosia: A case of reduced attentional "spotlight"? *Cortex, 28,* 601–621.

Tyler, R. (1968). Abnormalities of perception with defective eye movements (Balint's syndrome). *Cortex, 4,* 154–171.