People are unable to recognize or report on their own eye movements

Alasdair D. F. Clarke
School of Psychology, University of Aberdeen, Aberdeen, UK

Aoife Mahon
School of Psychology, University of Aberdeen, Aberdeen, UK

Alex Irvine
Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

Amelia R. Hunt
School of Psychology, University of Aberdeen, Aberdeen, UK

Abstract

Each eye movement we make brings new information into our visual system. The selection of each fixation is the result of a complex interplay of image features, current task goals, and biases in motor control and perception. To what extent are we aware of the selection of eye movements and their visual consequences? Here we use a converging methods approach to answer this question in three diverse experiments. In Experiment 1, participants were directed to find a target in a complex scene by a verbal description of it. We then presented the path the eyes took to find the target together with those of another participant. Participants could only identify their own path when the comparison scanpath was searching for a different target. In Experiment 2, participants viewed a scene for three seconds and then named some objects from the scene. When asked whether they had looked directly at a given object, participants' responses were primarily determined by whether or not the object had been named, and not by whether it had been fixated. In Experiment 3, participants executed eye movements towards single targets, and then viewed an animated replay of either the eye movement they just executed, or that of someone else. Participants were at chance to identify their own eye movement, even when it contained large under- and overshoot corrections. The consistent inability to report on oneâĂŹs own eye movements across experiments suggests awareness of eye movements is extremely impoverished or altogether absent. This is surprising given that information about prior eye movements is clearly used during visual search, motor error correction, and learning.

Although it has long been known that attention can be deployed to locations in a scene in the absence of eye movements (known as "covert attention", e.g. Posner (1980)), many experimenters have concluded that it is not possible to execute an eye movement to a location without also attending there (e.g. Hoffman & Subramaniam, 1995; Shepherd, Findlay, & Hockey, 1986). Indeed, the oculomotor readiness hypothesis (Klein, 1980), also known as the premotor theory of attention (Rizzolatti, Riggio, Dascola, & Umiltá, 1987) suggests that covert attention and eye movements are co-dependent; covert spatial attention is simply an eye movement that is prepared but not executed, making eye movements an "extreme" form of spatial attention. Under this idea, one might expect the process of attentional competition, selection, execution of the saccade, and accrual of new information arriving at the fovea to leave a lasting imprint on conscious experience. Eye movements also generate a massive signal in visual cortex. Visual areas are mapped retinotopically, such that every eye movement we make changes the contents of all the receptive fields of visual neurones from the retina all the way to higher-order visual areas (for a review, see Wurtz, 2008)). Given the enormous perceptual changes induced by eye movements, together with the fact that these changes are self-generated and have been suggested to necessarily involve attention, it seems reasonable to assume we should be able to report on these changes with some degree of accuracy.

Two relatively recent studies have concluded that awareness of eye movements is present, but limited. Foulsham and Kingstone (2013) presented participants with a series of photographs of scenes to remember. After completing a memory task, two versions of each scene were presented simultaneously. Overlaid on one image was an array of dots depicting the locations the observer had fixated while viewing that image during the initial display phase, and on the other image were fixations from someone else viewing the same image. Despite the approximately ten minute delay between viewing and recognition, accuracy was above chance (but only slightly, at 55%). Marti, Bayet, and Dehaene (2015) examined whether individuals can successfully reproduce the sequences of saccades they just executed on single trials of a visual search task using a computer mouse to click on the spatial locations in the order in which they looked at them. The spatial similarity of the series of mouse-clicks to the immediately preceding sequence of saccades was higher than for a saccade sequence produced in response to the same search array presented earlier in the experiment, suggesting some memory for the fixations that were just generated. However, participants missed many fixations and also reported many fixations that did not occur. The authors concluded that participants' ability to recreate their eye movements was limited, and may represent introspection of covert shifts of attention as opposed to eye movements per se.

Conscious experience is a notoriously difficult subject of study. In the case of reporting on one's own eye movements, there are a number of challenges to overcome. One of these challenges is that, when asked to identify or reproduce one's own fixation patterns, there are multiple ways to achieve above-chance performance. The Foulsham and Kingstone (2013) experiment used natural scenes, which participants were instructed to try and remember. Discrimination of one's own fixations compared to someone else's on the same scene could indeed be driven by a memory of the experience of looking at a particular object, as the authors suggest. But accuracy could also be driven by a memory of a particular object simply having been in the scene, leading to the inference that if they remember it, they probably looked at it (which, as we will discuss later, is usually true). In other words, accuracy to identify one's own fixations could be inflated by a memory of the existence of particular objects in a scene. On the other hand, fixations were presented to the participants with no timing or order information, and the presentation took place following a long delay relative to the initial viewing. The minimal information and delay may lead to an underestimation of how well people can recognize their own eye movements. The Marti et al. (2015) experiment overcame this latter limitation by requiring participants to report their fixations immediately following each search. While this removes the delay and may improve memory, it also introduces the potential for participants to adopt a strategy of generating more easily identifiable or reproducible fixation sequences during the initial search. Although the authors attempted to rule this out by comparing search during the introspection phase to an initial block of search only (with no introspection), small changes in strategy would be very difficult to detect and could be sufficient to elevate accuracy above baseline. More generally, it is difficult to rule out the possibility that a strategy, intention, or task can be recognised from a given eye movement or scanpath (e.g. Borji & Itti, 2014). Participants may be able to infer mental state from some scanpaths and compare this to their own remembered mental state while they viewed the same scene. This is not the same as actually remembering where they looked.

Although both of these studies suggest some ability to introspect and report on one's own eye movement, earlier work indirectly suggested participants are not aware of the perceptual effect of their own eye movements. In the double-step saccade paradigm, participants move their eyes towards a peripheral target which can jump to a new location during the movement. Participants tend to make a corrective saccade to the new target position, even though they report little to no awareness of the target's movement itself (Becker & Jürgens, 1979; Bridgeman, Hendry, & Stark, 1975). The focus in this area of research has been on awareness of information presented during saccades, rather than awareness of one's own movements, however, so it is not known whether and to what extent participants may be able to report on their corrective movements when explicitly asked about them. In other visually-guided actions such as reaching to a target, participants are able to immediately reproduce a reaching trajectory with an error induced by a target jumping to a new position, even when they are unaware of the jump itself (Johnson & Haggard, 2005; Johnson, van Beers, & Haggard, 2002). This suggests there is a stored memory of the motor correction in the absence of awareness of the signal that elicited the correction in the first place. It makes sense for motor processes to store and use information about motor corrections in order for motor learning to take place; indeed, after repeated trials in which the motor target moves to the same new location, both eye (McLaughlin, 1967) and hand (de Graaf, Pélisson, Prablanc, & Goffart, 1995) movements will be directed towards the expected final position of the target rather than its initial position, even though participants remain unaware of the target position shift. The motor system may be able to store and use information in the absence of explicit perceptual awareness of that information (e.g. as suggested by the dual-route model of visual perception, (Goodale & Milner, 1992)), so the existence of a correction does not necessarily mean participants will be able to report on it.

Further results suggest saccade errors can go unnoticed even when they are quite large. Mokler and Fischer (1999) found self-reporting of saccade errors to be relatively poor. In their experiment, a peripheral target would appear and the correct response was to generate an eye movement in the opposite direction (known as an antisaccade). They found an average error rate of nearly 20% with the majority of these errors consisting of an initial saccade towards the target followed by a corrective saccade. Participants failed to recognise around 50% of these errors **(see also Robinson and Irwin (2016))**. Similarly, in oculomotor capture (e.g. Theeuwes, Kramer, Hahn, & Irwin, 1998) the participants' task is to move their eyes to a color singleton. On some trials a sudden onset is added to the display at the same time as the colour singleton is revealed. On a substantial proportion of trials (30-80%, depending on timing and distance conditions) the eyes are directed towards the sudden onset, even though it not relevant to the task. The authors "*explicitly asked participants whether they were aware that the appearance of the new object affected their eye movements. Observers indicated that they were sure that their eye movements were not affected by the appearance of the new object.*" This is consistent with the persistence of erroneous saccades towards the sudden onset in both of the above studies, given that participants would likely correct their error if they knew they were repeatedly making it. Inconsistent, however, are the results from Belopolsky, Kramer, and Theeuwes (2008), who conducted an ERP study directly investigating awareness of errors in an oculomotor capture experiment. After each trial participants were asked if they moved their eyes directly to the target or not. Participants accurately reported the error on nearly two thirds of error trials. However, the proportion of trials on which the eyes were directed to the sudden onset was comparatively low in this study (16%), likely because onsets were constrained to locations that were far from the target position. This likely made erroneous eye movements towards onsets less frequent, and more noticeable, than in the original study.

The Theeuwes et al. (1998) and Mokler and Fischer (1999) experiments had the advantage of using very simple geometric stimuli and very constrained saccade task parameters which limit the potential for alternative strategies for reporting on one's own eye movements. **However, they do rely on self-report on eye movement accuracy, explicitly reported after every trial. This leads to the concern noted previously that participants may deliberately alter the way they move their eyes to make it easier to report where they looked, suggesting these studies may be overestimating awareness. On the other hand, exposing participants to repeated trials in which single eye movements to simple stimuli are required could impede the ability to discriminate one eye movement from another. Under more natural circumstances, in which eye movements are employed in the service of gathering information in a more natural scene, we may find people are reasonably good at reporting on the consequence of the process of competition and selection that drives eye movements to particular objects or locations.**

It should be clear from the above discussion that the extent to which we are aware of our own eye movements is an important but difficult question to answer. No single approach can satisfactorily address all the challenges associated with probing self-awareness of eye movements. Our goal in this series of three experiments is therefore to use converging methods to understand whether, and to what extent, participants can recognize their own eye movements. In Experiments 1 and 2, we use natural and complex scenes and sequences of multiple fixations, and participants had to explicitly recall their fixations, similar to the approach used by Foulsham and Kingstone (2013) and Marti et al. (2015). However, we also manipulate the availability of alternative strategies for elevating accuracy above chance, namely using inferences about task or search goal (Experiment 1) or memory for objects in a scene (Experiment 2). We find that in the absence of these strategies, accuracy to recognise or report on one's own fixations is minimal and does not differ from chance. In Experiment 3, participants made simple saccades to single targets and had to state whether the animation of the eye movement following the trial was their own or someone else's. We perturbed the target position on a proportion of trials to induce corrective saccades. This allowed us to determine whether or not participants were able to recognise their own saccades when they contained corrections, or to correctly reject saccades that contained corrections when their own saccades were accurate. Participants had a bias to think accurate saccades were their own, and were otherwise at chance at this task. The results from these three diverse experiments converge on the same conclusion: people are generally not directly aware of their own eye movements, but they have many strategies at their disposal that they can use to elevate their accuracy to report on their eye movements above chance.

General Methods

Set-up

All experiments conducted within this study were undertaken in the Eye Movements and Attention laboratory at the University of Aberdeen. Equipment and methods that are common to all the experiments are described here. Additional materials are outlined under each experiment methods section. Experimental scripts were created and run using MatLab with the PsychToolBox (Brainard, 1997) and Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002). A PowerMac running OSX 10.8.2 was used and stimuli were presented on a Sony Trimaster EL computer screen at a resolution of 1080 × 1920. Participant responses were recorded using an Apple keyboard with numeric keyboard, a mouse or voice recording applications, depending on the experimental design. An EyeLink 1000 (SR Research, Mississauga, Canada) was used to track eye movements. The protocol for each of the experiments was reviewed and approved by the Psychology Ethics Committee at the University of Aberdeen.

Participants

Participants were recruited from the population of students and other members of the academic community at the University of Aberdeen. All participants had normal or corrected-to-normal vision. The experiment was conducted with the full understanding and signed consent of each participant. Participants were remunerated £5-10 for their time, depending on the length of the experiment. Some participants took part in multiple experiments.

Analysis

We have chosen to follow recent advice from Cumming (2013) on the reporting of results in psychology research. Namely, we will avoid using *p*-values and null-hypothesis significance testing wherever possible. We expect to find a range of abilities for different observers, and our aim is to measure, report, and interpret that range under different tasks and conditions. Where appropriate, we use general linear mixed-effect models (from the `lme4` package for **R**) to estimate effect sizes and standard errors while factoring our random effects associated with differences between individual observers and images. 95% confidence intervals will be obtained by bootstrapping using the `confint` function. The data from these experiments has been made publicly available[1].

Experiment 1: Visual Search

In our first experiment, we were interested in discovering whether participants are able to recognize their own scanpath relative to someone else's viewing the same image. **We followed the methods of Foulsham and Kingstone (2013), who observed accuracy at this task that was only just above chance (55%). This result suggests participants are able to do this successfully for a few images or scanpaths, but are at chance for the majority of trials.** On the one hand, it is possible that participants are using a memory for which objects were fixated to do the fixation recognition task, as the authors suggest. However, an alternative route through which participants could achieve above-baseline accuracy in the scanpath recognition task is to infer a particular strategy or goal from the scanpaths they are shown and then match this to their own remembered goal or strategy. This strategy could also pose an alternative explanation for why Marti et al. (2015) found that participants could reproduce their own scanpaths with a mouse immediately following search: a search path could be reproduced by remembering and implementing a strategy or mental state evoked during the initial search, rather than remembering the eye movements per se. This experiment tests this alternative explanation by manipulating search instructions.

We asked participants to search for a particular target in cartoon images taken from the Where's Wally children books. Participants then had to discriminate their own scanpath from that of someone else. Following Foulsham and Kingstone (2013), participants did not carry out the scanpath discrimination task until after all images had been viewed, so they could not strategically change their search behaviour to make their own scanpaths more recognizable. A description of what the participant needed to find in each scene was provided verbally over headphones at the start of each trial. We manipulated the search instructions to control how the participant searched, and therefore how similar a given participant's fixations were likely to be relative to those of other observers. Specifically, relative to the comparison scanpath, participants were given a) the same search target b) a different search target, c) a description of the search target that first provided a salient landmark and then the target (i.e., "near [the landmark] is [the target]") and d) a description in the opposite order (i.e. "[the target] near [the landmark]"). We expected that scanpaths that came from searches prompted by the same target description should be more difficult to discriminate than those prompted by different descriptions of the same target, with search for a different target altogether providing the most easily discriminated scanpath.

---

[1] insert URL here if paper is accepted

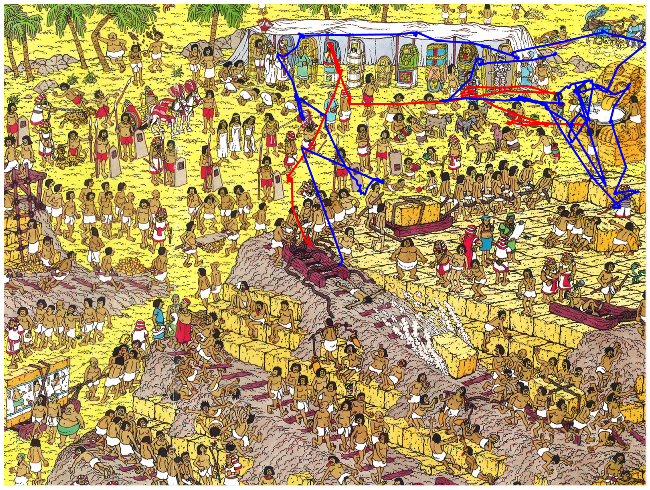


Figure 1. Example stimulus with scan-paths overlaid. The four referring expressions for this trial were (A) "at the upper right, the sphinx"; (B) "at the upper right, the man holding the red vase with a stripe"; (C) "at the upper right, the man holding the red vase with a stripe to the left of the sphinx"; (D) "at the upper right, to the left of the sphinx, the man holding the red vase with a stripe on it".

Methods

Participants. Thirty-two participants (median age 23, range = 19 - 42 years old, 21 females) took part in the current study. One participant was dropped due to excessively long reaction times. Due to the design of our study (where each participant's scanpath is compared to that from the previous participant), the participant that followed this rejected participant also had to be discarded from analysis as they would have been comparing their typical length scanpaths to excessively long scanpaths.

Stimuli. Stimuli (images and search instructions) were taken from Clarke, Elsner, and Rohde (2013b), in which participants were given a target and asked to provide a description of how to find it (known in linguistics as a referring expression). We used as a target one of the sixteen targets per image used in this previous study, as well as one of the landmarks that had been spontaneously named by participants to help others to locate the target. The following variations were constructed to give us four search instruction conditions:

(A) Find the landmark

(B) Find the target

(C) Find the target next to the landmark

(D) Next to the landmark, you will find the target

Search instructions for each image all began with the same regional descriptor ("*in the upper left...*"), followed by the specified search target. Twenty-eight images were used in total. An example image with accompanying search instructions for the four conditions is shown in in Figure 1.

Procedure. Immediately following image onset, an audio recording of the search instruction was played to participants over headphones, giving them the necessary information required to find the target. Participants pressed the space bar on the keyboard when they had found the specified target. They were then required to use the mouse to click on the target. This was done so that we had a record of search accuracy and participants were not able to just press space without finding the target. Only eye movements from image onset to the space bar response were used.

After completing all 28 visual search trials, participants then carried out the second section of the experiment. This consisted of viewing all 28 images again (in a new random order), with two scanpaths (in red and blue) drawn on top, as illustrated in Figure 1. Scanpaths were represented as a series of line segments connecting the gaze points recorded by the EyeLink. Gaze samples falling outside of the image **were** discarded. One scanpath illustrated the participant's eye movements from searching the same scene earlier in the experiment, while the other scanpath showed the behaviour of the previous participant in the experiment (searching the same scene, but possibly with a different search instruction). Participants were asked to decide, by indicating the colour, which of the two scanpaths was their own (hence the first participant did not carry out this part of the experiment). Participants were not informed of the task in this part of the experiment until after they had carried out the first part, thus excluding the possibility that participants could modify their search behaviour to make it easier for them to remember their own eye movements.

Discarded Trials. A number of trials were excluded from analysis and due to the design of the study (with each participant comparing their performance to that of the previous participant), if a trial was excluded then we had exclude the subsequent participant's corresponding trial. Trials with very short (<1s, 11 trials) or long (>60s, 1 trial) reaction times were excluded. We also excluded trials with a delay of >5 seconds between the time when participants pressed the space key to indicate that they had found the target, and when they clicked on the target with the mouse to verify they had found it (6% of the remaining data points). Unlike in reaction times, there was no clear cutoff to use to define outliers (there is a long tail of click times with a maximum of 20 seconds) and so we ran the analysis with and without this exclusion criterion. Results were similar in both cases.

Results

The difference between the conditions in terms of reaction time is **summarised in Figure 2. A complete analysis of these data is beyond the scope of this paper, and can be found in Clarke, Elsner, and Rohde (2015).**

Figure 3 shows the accuracy to discriminate one's own scanpath from that of another person for all possible pairs of search instructions, averaged across participants and images. Mirrored conditions (e.g. accuracy for the comparison of landmark to target and target to landmark) were averaged. It is clear from this figure that participants are generally at chance to correctly identify their own eye movements when the reference scanpath is from a person performing the same task (all points shown inside the grey box). The only conditions at which performance differs from chance was when the participant was searching for a different target from the reference (A-B, (77% with a 95% confidence interval of $[66\%, 86\%]$), A-C, (64%, $[52\%, 75\%]$) and A-D (66%, $[54\%, 76\%]$): in these three comparisons the search for the landmark alone (A) is compared to search for the target, either with or without the

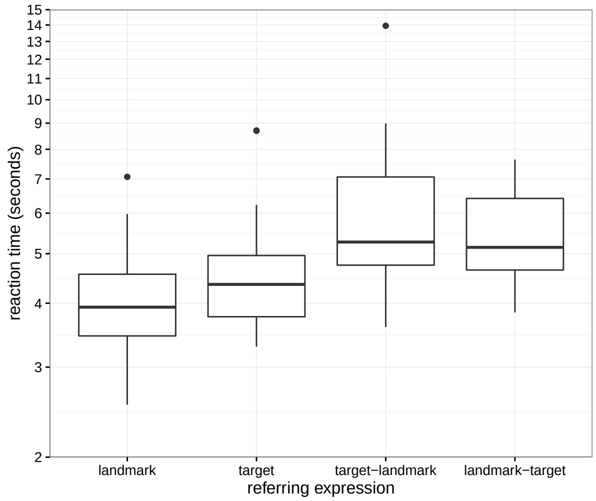


Figure 2. Boxplot showing distribution of each participant's median reaction times for each condition. **The y axis is on a log scale.**

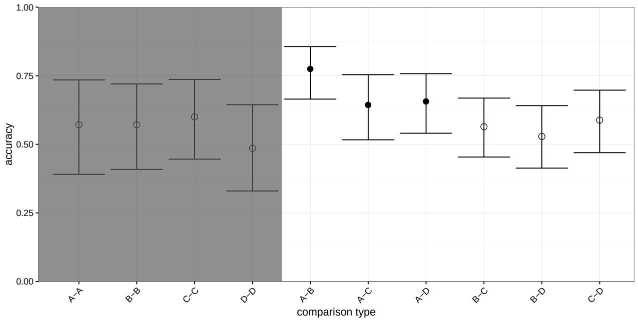


Figure 3. Accuracy to discriminate one's own scanpath relative to a reference scanpath from the same (grey background) or a different (white background) search instruction condition. Error bars show 95% binomial confidence intervals. Filled dots indicate conditions that are above chance based on whether or not these intervals include 50%. Letters denote the search instruction conditions: A=landmark; B=target; C=landmark-target; D=target-landmark.

landmark given as a part of the target description). We also analysed these results more completely, taking participant and image effects into account using a linear mixed effects model. The conclusions we drew from this analysis were essentially the same as those that can be drawn from simply examining Figure 3. We have included the model fits in Appendix A.

The results from this experiment suggest that participants are not very good at discriminating their own eye movements from another person's. **Participants were above chance to identify which scan path was their own only when the other person was searching for a different object. This may be because the two scanpaths tended to be objectively less similar when they were searching for a different object, however, visual inspection of the scanpaths suggests that there were many trials in which participants identified the wrong scanpath as their own, despite salient differences between the two. (See Appendix).** This suggests it is not a memory of eye movements per se that participants are using to correctly select their own scanpath, but a more indirect inference about the target of search compared to what they remember having looked for. However, this experiment included a complicated search and a long delay between when the search occurred and when the participant was asked to report on it; asking participants a simpler and more immediate question about where they recently looked may afford better performance.

Experiment 2: Objects in Scenes

As discussed in the introduction, another way in which participants could perform above chance in the scanpath recognition task is by assuming the objects they remember seeing in the scene are objects that they themselves fixated. As we will demonstrate, this is a strategy that can elevate participants above chance, because it is in fact the case that objects that are fixated in a scene are also more likely to be reported. Our second experiment explicitly tested whether people used this recognition strategy by testing memory for objects in the scene together with memory for which objects were fixated.

A previous study by Clarke, Coco, and Keller (2013a) asked participants to verbally report objects from a natural scene immediately after presentation. They found the objects that were named tended to also have been fixated but there were also many objects that were named but not fixated, and objects that were fixated but not named. We took advantage of this existing dataset to predefine a series of objects in the set of natural scenes used by Clarke et al. (2013a) that were likely to be fixated and/or named. We asked participants to view a series of images. After each image was removed, they were asked to name objects they remembered from the scene. On half of trials, they were then also asked if they had "*looked directly at*" a specified query object or not. We expected participants would be able to correctly confirm that they looked directly at an object when they had also named it. We were particularly interested in whether participants could also correctly state (i) that they did not fixate an object that they just named as having been in the scene, and (ii) that they did fixate an object even though they did not name it. Being able to accurately classify objects as having been fixated or not fixated irrespective of whether or not they named them as having been in the scene would be a clear indication that people are able to remember their fixations separately from the objects in the scene.

Methods

Participants. Thirty-two participants (median age 24.5, range = 20 - 62 years old, 22 females) volunteered for the current study. One participant was dropped due to a recording error.

Stimuli. Stimuli were taken from the set of annotated images used by Clarke et al. (2013a). In the original study, 24 participants viewed each image for three seconds while wearing an Eyelink II eye tracker. After each scene was removed from display, participants were asked to name objects they could remember from the scene (usually about five). Using these data, we searched the set of 100 images for objects that matched the following criteria:

- Size of object was between 1000 and 60,000 pixels (1/8th of image size)
- Object was unique in image (for example, the only chair in the image)

(a) “people, bench, geese, grass” (b) “laptop, mouse, chair, bed”

Figure 4. Two example trials. In the first image, the query object was the bottle (on the left by the bench) and the participant correctly answered that they had not fixated it. In the second image, the query object was the laptop and we can see that while the participant named it, they did not directly fixate it during the trial.

- Objects that were:
  - Named by at least 70% of participants and fixated by at least 60%
  - Named by at least 70% of participants and fixated by at most 33%
  - Named by at most 33% of participants and fixated by at least 60%
  - Named by at most 33% of participants and fixated by at most 33%

We created seven trials for each condition, giving a total of 28 critical trials. Note that this is an a priori estimate of the actual number of trials that were expected to fall in each condition based on previous data; the actual number of trials falling into each of the conditions above was determined posthoc based on the data collected from a new set of participants. A further 28 filler trials, selected from the same dataset of photographs, were added to give a total of 56 trials. The filler trials were included in an attempt to decrease the amount of attention the participants would give the fixation recognition question during scene viewing, and instead focus on the memory question that was asked after every trial. Two example images are showing in Figure 4.

Procedure. Images were shown in a randomised order. Before each image, participants pressed the spacebar while fixating a point in the centre of the screen and a drift correction was carried out. The image was then presented for three seconds before being removed from screen. Participants were asked to name aloud as many objects as possible from the scene. In order to avoid them just naming the most salient object, or using long term semantic memory to make a long list of educated guesses, they were encouraged to name “around five.” Responses were audio recorded for later transcription. Participants pressed the space key when done. If the trial was one of the 28 critical trials, they were then asked:

“Did you look directly at the `query object`?”

Table 1

Number of Trials in each Condition.

Fixated	Named	Number of Trials	Median Num. of Trials per Observer
Yes	Yes	232	9
Yes	No	296	8
No	Yes	96	3
No	No	232	8

Participants responded by pressing the `y` or `n` key on the keyboard. There was no time limit.

Analysis. This is a binary response task with an unbalanced design. As such, a simple accuracy measure such as percent correct is not suitable for characterising performance, because individual participants may have a response bias that could run the same or the opposite way of the bias in what is the correct response. For example, if there is a general bias to say "yes" to the question of whether or not a given object was fixated, the proportion correct for a given individual in their reports will depend on the extent to which they happened to look at the queried objects in the set (rather than on their ability to remember looking at it, which is what we are actually interested in). We therefore need an appropriate measure of how well people can discriminate their accurate from their inaccurate eye movements. Hence we will present our results using two statistics commonly used in the classification literature: precision and recall. If we are trying to classify A from B, then the definitions are as follows:

- Accuracy: the proportion of all items successfully classified.
- Precision: the proportion items classified as A that are actually A.
- Recall: the proportion of items belonging to class A that are classified as A.

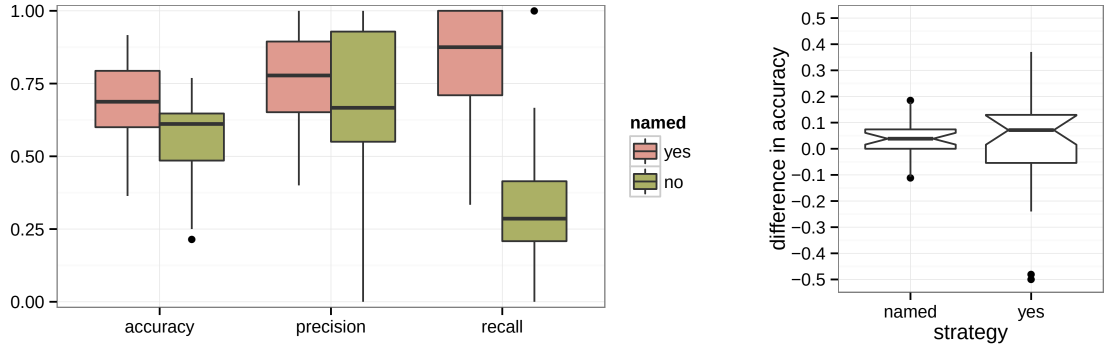
An object was considered to have been fixated by a participant if at least one of their fixations fell within a polygon marking the outline of the object.

Results

Participants appeared to manage the naming task reasonably well, naming at least four objects on nearly all trials. Two example trials with overlaid scanpaths and named objects are shown in Figure 4.

As explained above, image-object pairs were selected from an existing dataset to try and obtain an approximately equal number of trials containing query objects of roughly equal size in each of four conditions (fixated/not fixated × named/not named). However, as we have no control over which objects observers look at or name, we expected this to vary from person to person. The trials were therefore re-categorised into the same four conditions using the data based on our new participants' actual behaviour. The number of trials in each condition is shown in Table 1. With the exception of not fixated yet named objects, we have an approximately even split of trials over condition.

We then examined accuracy in the fixation recollection task as a function of whether or not the queried object was named. The results are shown in Figure 5a. Observers

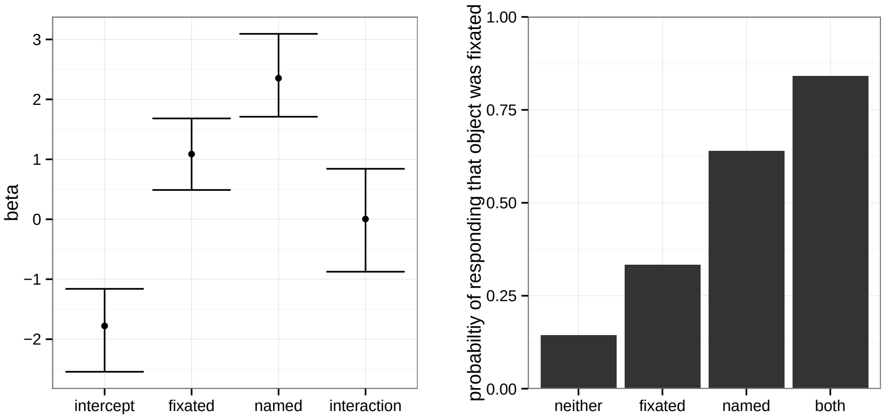


(a) (b)

Figure 5. (a) Accuracy statistics for the objects experiment. Accuracy to state whether an object has been fixated or not is modestly above chance. However, for objects that were not named, participants tend to report not having fixated it even when they had, resulting in a very low recall score. (b) Difference scores when comparing the observed accuracy to expected accuracy under a strategy of (i) responding that you fixated the object if and only if you named it ("named") and (ii) simply responding that you fixated every query object ("yes").

are more accurate at determining whether or not they fixated named objects on all three measures (accuracy: median of 69% for named objects compared to 61% for unnamed objects; precision: 78% compared to 67%; and recall: 88% compared to 24%). The largest difference in between conditions is in recall, where the recall of fixated unnamed objects is very low. This means that participants tended to incorrectly state that they did not fixate an object when it was an object they had not named. When we compare these results to some simplistic baseline strategies (Figure 5b), we can see that human performance only marginally outperforms some very simple response strategies: in particular, performance is quite close to what would be expected if participants simply stated they fixated those objects which they had named (median difference of 3.8%), with about 25% of participants under-performing this strategy. Even more participants under-performed the strategy of simply stating "yes" every time they were asked if they fixated a particular object.

We can further analyse these data by fitting a general linear mixed model to investigate how the participant's response (whether they believe they have looked at the query object or not) is influenced by (i) whether they have actually looked at the object and (ii) whether they named the object. We fit a model with random intercepts for both participant and image and the fixed effects are shown in Figure 6. The results demonstrate that although actually fixating the object does influence the participant's response (the probability of participant responding that they had fixated an object which they had neither fixated or named was 14%, while for objects which were fixated but not named this rises to 33%), the influence of whether or not a given object was named is even larger (63%). Following advice on areas-of-interest padding (Orquin, Ashby, & Clarke, 2015) we re-analysed the data, expanding the radii of the AOIs between 1% and 25%. We found that this made little difference to the model, and all estimated coefficients remained within the 95% confidence



(a) (b)

Figure 6. (a) Fixed effects from general linear mixed model with 95% confidence intervals. (b) Logistic transform of fixed effects for each experimental condition. We can see that a named object is more likely to be reported as fixated than one that actually was fixated.

intervals presented in Figure 6.

We also explore the extent to which total fixation duration has on the likelihood of a participant reporting that they had fixated an object. We analyse the subset of the data consisting of all trials in which the participant fixated the target object, and rerun the general linear mixed model replacing the binary `fixated` factor with log total dwell time. Bootstrapped 95% confidence intervals suggest some evidence of a possible relationship between fixation duration and participant response ($\beta = 0.6$, CI=$[0.065, 1.333]$). However, these results should be treated with caution given the small sample size and documented relationship between fixation duration and the likelihood of an object being named (Clarke et al., 2013a)

Our results suggest that observers have little to no memory for what they fixated, and simply respond that they looked at the objects that they named. This is in contrast to the conclusions of Foulsham and Kingstone (2013), but we believe participants could have employed an object-memory strategy in their experiment at least on some trials, which would be sufficient to boost accuracy above chance. An important design difference between our study and Foulsham & Kingstone's is that they did not ask participants about their eye movements until all images had been viewed (as in our Experiment 1). Here we asked participants about their eye movements directly after they had viewed the image, which should, if anything, have improved accuracy relative to a delayed recollection. Asking immediately following the trial does, however, mean that over the course of the experiment, participants would potentially be able to alter their viewing strategy to increase their performance in the fixation recollection task. However, given the low accuracy at this task, this does not appear to have happened.

Experiment 3: Double-step Saccades

The last of our three experiments investigated whether individuals could identify their eye movements in a simple saccade task towards a single visual target, in which participants were fully aware that they would need to report on their eye movements on every trial. Participants simply made a saccade to a peripheral target. On half of trials, the peripheral target jumped to a new position while the saccade was being executed. Participants tend to perform a corrective saccade to the target's new position (known as a double-step saccade), although they typically do not notice the change in target position (Becker & Jürgens, 1979; Bridgeman et al., 1975). The disruption to the saccade introduced by the position shift in the saccade target could, however, provide a signal to the participant that they could use to differentiate their own eye movement from someone else's, if this is indeed possible. After each trial, participants viewed an "instant re-play" animation depicting either the eye movement(s) the participant just executed or those of someone else. An animation was used to give the participants as much temporal and spatial information about the movement as possible. The aim of the experiment was to determine if participants were sensitive to executing either single or double step saccades and as a result if they can discriminate their eye movements from someone else's.

Methods

Participants. Twenty-eight participants (17 female) volunteered for this experiment, aged between 21 and 29 years old.

Stimuli. Stimuli consisted of a $2° \times 2°$ fixation cross displayed upon a blank screen at the start of each trial. Subsequent single or double fixation crosses were presented either right or left of the original central fixation cross depending on condition.

Procedure. Participants were given a practice block of 16 trials to allow them to increase the speed at which they made their eye movements as well as improve their final fixation accuracy so they landed at the centre of the movement target. Participants had to successfully complete a minimum of 12 of these trials in order to carry on to the complete the full experiment (criteria for "successful" is given below); if they failed to complete 12, they repeated the practice until they reached this criterion. The experiment consisted of six blocks of 37 trials, however, trials in which participants made an incorrect movement or were above the speed threshold were omitted. At the start of each trial a fixation cross was displayed at the centre of the screen. In order to begin each trial participants fixated the cross and pressed the space bar. After a successful drift correction a smaller central fixation cross was displayed and at the same time an additional second movement target cross was presented at one of three locations to the left or right of the central fixation cross (six locations in total). Participants were instructed to simply move their eyes to the target as quickly and accurately as possible. On half of the trials the fixation point remained in the same location throughout the trial. On half of these trials the second cross was 11.5° from the centre, and the other half of trials the cross was 15.8° from the centre. In the remaining half of the trials **where** the cross was initially shown at 13.7° from the centre, and either jumped inwards (to 11.5°) or outwards (to 15.8°) once a saccade was detected (simply defined as when the x-coordinate of the current gaze location was more than 1.4° from the centre of display).

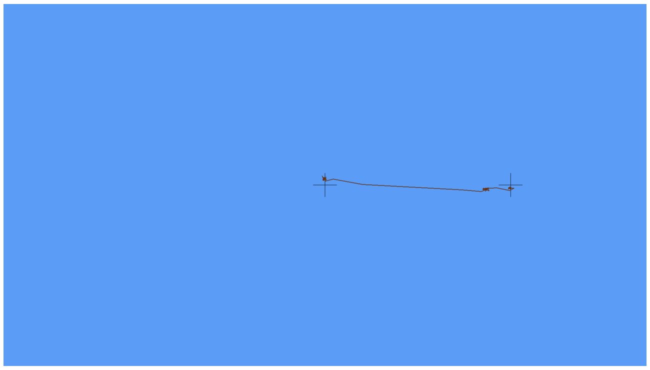


Figure 7. An example of a gaze replay presented to participants. In the actual experiment, this line was traced out in real time.

To successfully complete a trial, participants were required to fixate the movement target (defined as a period of 50ms in which all gaze samples were within 2° of the centre of the fixation cross). If this criteria was not met within 600ms of target onset, the trial was counted as invalid and the participant was shown a red screen for 3000ms. If participants successfully completed a trial, a further 50ms of gaze samples were collected, followed by a blank screen for 500ms, after which participants were shown a replay of gaze behaviour during the trial. In half of the trials, they were shown their own eye movements in an animated replay, drawn on the screen, sample-by-sample, in real time. In the other half of the trials, they were shown a pre-recorded eye movement collected in a pilot experiment. The replay displayed two fixation crosses: the central fixation cross and the movement target cross at either the final inwards or outwards position (see Figure 7). Participants were asked to identify whether the gaze replay shown was theirs or someone else's. If they believed the gaze replay was theirs they pressed `y` on the keyboard, if they believed the gaze replay shown wasn't theirs they pressed `n`. They were informed that the replay would be their own eye movement 50% of the time.

The colour of the display screen changed colour depending on what phase of the trial the participant was completing. This was implemented to help participants keep track of the task. In the first phase, in which they were making, or preparing to make, an eye movement, the screen was presented as pale orange. In the second phase, when participants were watching, or preparing to watch, the gaze replay, the screen was blue. See Figure 7 for an example.

Discarded Trials. Any trials in which participants did not fixate the movement target for a minimum period of 50ms, in which all gaze samples were within 2° of the centre of the fixation cross, were excluded from further analysis. Also any trials in which this criteria was not met within 600ms of target onset were counted as invalid and discarded from further analysis. Under these criteria, a mean of 73% of trials from each participant were included in the analysis (with a large range of individual differences: 40% - 94%). Four participants had over half of their trials discarded.

	Inwards target	Outwards target
forwards correction	122	759
accurate saccade	985	494
backwards correction	217	5

Table 2

Number of saccades falling into each category. ***Categories were defined with respect to the target's end position: trials in which the target moved inwards from the central position are grouped with the trials in which the target appeared, and remained, at the closer (to the centre of the display) position. Similar for targets that moved outwards, away from the central fixation cross.***

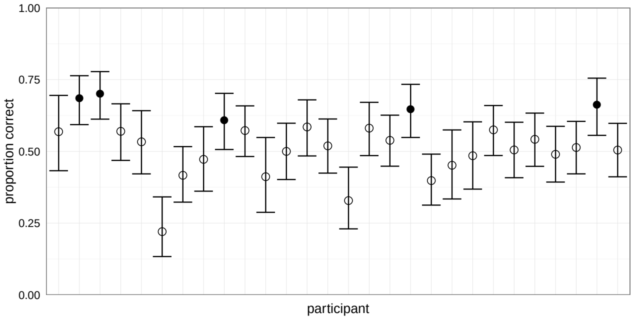


Figure 8. Accuracy with 95% Confidence interval for each participant. The filled dots indicate the participants who did better than chance.

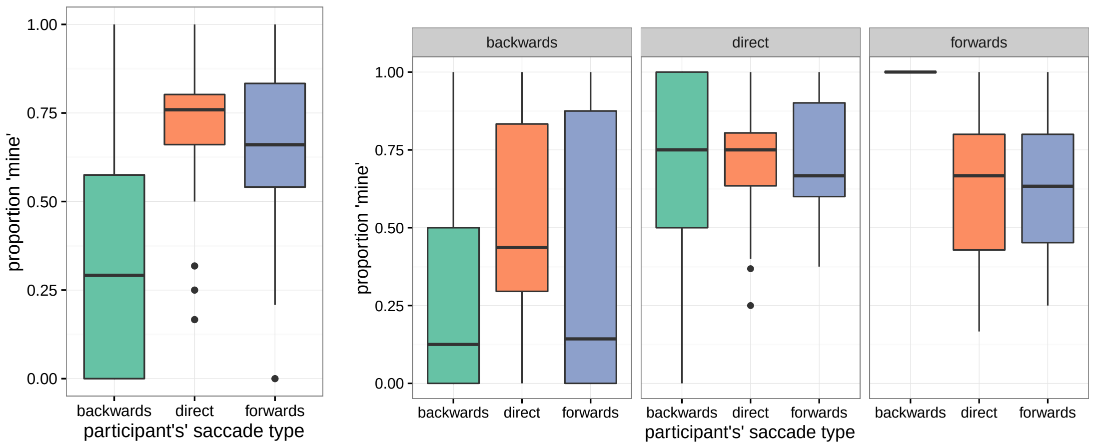
Results

Table 2 shows the distribution of saccade types by target position. Note, this collapses over trials in which the target appeared at the inwards/outwards position, and those in which it appeared in the central position and then moved.

Participants were required to accurately identify if the gaze path presented at the end of each trial was their own or someone else's. Accuracy to do so (in terms of the proportion of trials with a correct response) was poor (mean participant accuracy = 52%, with only five out of 28 participants managing to perform better than chance,with the most accurate participant achieving a score of 70%. Three participants performed below chance.

Although overall accuracy was around 50%, classification differed between comparison conditions. In particular, when participants were shown a direct saccade or a saccade with a forward bias, they had a strong tendency to classify that saccade as their own. Conversely, participants rarely classified saccades with backward corrections as their own. As can be seen in Table 2, participants did perform far fewer backward corrections than forward corrections or direct saccades. This is clear from Figure 9a, which shows the tendency to "claim" (i.e. identify as "mine") saccades on trials in which the participant was shown their own saccade: when participants actually made a backward correction, they were far less likely to claim the saccade than if they made a forward correction or direct saccade.

We can also break the results down into similar categories on trials in which the participants were shown someone else's eye movements, shown in Figure 9b. Again, participants are more likely to (now erroneously) state a saccade is their own when it is direct



(a) (b)

Figure 9. (a) Proportion of trials on which participants stated the saccade was their own for only those trials in which they were actually shown their own saccade. When the saccade was direct or had a forward correction, they were far more likely to state it was their own than when it contained a backward correction. (b) Proportion of trials on which participants stated the saccade was their own, for trials in which they were shown someone else’s backward correction (first panel), direct saccade (second panel) or forward correction (third panel). The three bars within each panel separate the results based on which type of saccade the participant has themselves just executed; backward saccades were not very common, so this category should be interpreted with caution. In general, participants were just as likely to claim a saccade was their own when it matched the saccade they just executed as when it did not.

or forwards. Saccades that fell into a different category than the one that had just been executed were as likely to be claimed as saccades that were in the same category. This confirms that participants have a bias to identify correct saccades as their own, and incorrect saccades as having been generated by someone else. To the extent that participants do tend to produce correct saccades, this bias boosts their overall accuracy in the discrimination task, but it is clear they have little to no insight into their own saccade errors.

Discussion

Previous work has suggested that people have a surprisingly limited ability to report on their own eye movements. However, as explained in the introduction, the methods used to answer this question could have a large impact on the results. Therefore, by design, each of our three experiments were conducted using very different paradigms. Although each experiment tackles the question differently they all converge on the same conclusion that, in general, people are even less aware of their own eye movements than previous research suggested.

In Experiment 1 we varied which target the person searched for, and, more subtly, altered the linguistic structure of the target description to include a salient landmark description either before or after the target was specified. If participants can extract information about the goals or intentions driving the eye movements and compare it to their own, they should be better able to discriminate their own eye movements when compared against a person searching for a different target, and/or against a person searching based on different instructions from their own. Indeed, we found that participants' performance was highest (although still not particularly high) when they were comparing their scan path to one that was looking for a different target. However, when their scanpath was shown together with that of a different person searching for the same target, even when they were following a different search instruction, discrimination was at chance. This indirectly suggests that participants are basing their judgements on inferences about goals, rather than any detailed memory of their own actual eye movements. But could it just be too difficult a task to remember visual details well enough to discriminate them from a relatively similar foil shown up to ten minutes later? Probably not; experiments on recognition memory for small variations in object identity or position has shown this capacity to be surprisingly large, even when testing memory for thousands of similar objects at a delay of several hours (Brady et al., 2008). In the present study, participants were asked about their eye movements immediately following their execution in Experiments 2 and 3, and there was little evidence that this boosted performance much compared to Experiment 1, suggesting the poor performance is due not to fixational memory decaying over time, but to it having not been present in the first place. Another consideration is that we are concluding participants are "at chance" based on an aggregate score with confidence intervals that overlap chance (50%), leaving open the possibility that some of our participants genuinely perform above chance at identifying their own scanpaths. That said, it is possible that these participants (and/or the person with whom their scanpaths were paired) had features in their scanpaths that were unique and easily identifiable, such as systematic calibration errors or tremor. This could inflate accuracy above chance for some individuals, again because of indirect inferences as opposed to memory of one's own selection of fixations per se. Even though we cannot eliminate this possible alternative route to accurate performance, aggregate results are still at chance, suggesting at the very least that the majority of participants cannot accurately recall their own scanpath.

In Experiment 2 we established that individuals are able to use a memory of which objects were present in the scene as a reasonable approximation of which objects were fixated. Indeed, if participants simply report that they fixated any object they remember having seen, the overall accuracy of these reports would be about 60-70%, suggesting this is a good strategy for achieving above-chance accuracy even in the absence of any explicit memory of fixations themselves. The results when we split recognition performance by whether objects were named or not reveals that participants were relying on this strategy almost entirely, with very low recall for objects that were not named - in other words, participants thought they had fixated named objects even when they had not, and vice-versa. Indeed, having named an object was a better predictor of whether a given participant stated they had fixated it than whether or not they actually fixated it. That said, after accounting for the effect of naming, there was some variability in judgement of fixations that could be accounted for by the fixations themselves. This could suggest that participants had some memory for their own fixations independent of whether they remember the object in the scene. On the other hand, we asked participants to name "around five" objects in the scene, and it is quite plausible that some objects were encoded but then not subsequently named. This could be because participants could not easily put a verbal label on the object, or simply prioritised it less than other objects in the scene. In any case, those few trials on which participants report having fixated an object they did not name could represent trials on which they encoded, but did not name, the object. Given how infrequent these instances are, we can at the very least say that the dominant strategy for deciding whether an object was fixated or not is to use one's own memory of whether or not it was in the scene at all. Given that this strategy leads to above-chance accuracy in the fixation recall task, caution should be exercised while interpreting results from experiments in which participants can rely on memory for objects instead of fixations for accurate self-report (as they could in Foulsham and Kingstone, 2012). This self-report is not an indication of a memory for fixations, but rather a memory of what objects were or were not present in a scene.

In both Experiments 1 and 2 we find clear evidence that participants rely on alternative strategies rather than fixational memory when reporting on their own eye movements. This suggests we do not store a representation of fixations, at least not one we can easily report. However, in both studies the task was difficult, and the real purpose of the study was to some extent disguised; in the first experiment by not probing recognition memory until the end, and in the second by only asking about fixations on a smaller subset of trials. Based on these two experiments alone, the possibility remains that participants can in fact maintain a representation of their own eye movements as long as this information is task-relevant, but when this information is not consistently required it is not accessed, stored, or maintained. In the final experiment, therefore, we made it explicitly clear that the main task of the experiment was to remember and report on one's own saccades. The simpler saccade task circumvents the use of task inference or object memory, focuses on a memory of eye movements to a single target, and asks participants to immediately recognise an animation of this eye movement after every trial. We increased the variability of the eye movement by inducing corrections, and we used real-time animations to present as much spatial and temporal information to the observer as possible. Nonetheless, participants as a group were at chance at this task: only half of the trials participants identified as their own were actually their own. This conclusion is based on aggregate accuracy, however; there were five (of 28) participants who were significantly different from chance when their results are examined individually. As in Experiment 1, it may be that these participants do in fact have a clearer representation of their own eye movement trajectories than the other 23 participants, but there are at least two other possible explanations. One is that the eye movements of these five participants had distinctive characteristics (e.g. fixation tremor) that made their eye movements more easily detected relative to the comparison eye movement than the others. A second explanation is that these participants were simply better at noticing and using small systematic differences between eye movements to discriminate one group from another.

In all the above experiments, the experience of the eye movement itself is different from the medium we used to probe memory. In the first experiment, we showed scanpaths as a series of lines connecting fixated points, which provides spatial and sequence information, but temporal information was removed. In the second experiment the memory probe was even further abstracted from the actual experience of fixating the object because we asked verbally about an object that was present in a scene which was no longer visible. Even in the final experiment, in which we replayed the eye movement back in real-time, the experience of actually moving the eyes would have been quite different from the representation of that movement as a dot on the screen. We did consider making the animation an even closer match to the participant's own immediate experience of the eye movement they just executed by replaying the movement of the saccade target across the retina as a consequence of the eye movement. However, we assumed that naive participants would have a difficult time understanding what we were showing them, given that this is not the way people tend to report or describe their eye movements; most people intuitively refer to the image as remaining fixed and their eyes changing position on that image, rather than the eye movements changing the position of the image itself. The anecdotal observation that we do not easily or naturally comprehend the effect of our own eye movements on the image falling on our retina is broadly consistent with our conclusion that we do not have a very accurate representation (if any) of these movements in the first place. The lack of awareness of the visual impact of one's own eye movements is often referred to as saccadic suppression (Matin, 1974). The threshold to detect stimuli, particularly motion signals, is elevated during saccades, likely due to a combination of post-saccadic masking (Campbell & Wurtz, 1978; Ibbotson & Cloherty, 2009) and active suppression of some pathways and signals (Bremmer, Kubischik, Hoffmann, & Krekelberg, 2009; Ross, Burr, & Morrone, 1996). There is clear evidence of ongoing visual processing of form and flicker during saccades Hunt, Chapman, and Kingstone (2008), Watanabe, Noritake, Maeda, Tachi, and Nishida (2005), indicating that suppression is selective to some channels. As noted in the introduction, however, participants tend to be unaware of displacements in the saccade target position that occur during saccades (Bridgeman et al., 1975) despite the fact that rapid corrective saccades are executed to the new target position (Becker & Jürgens, 1979), and after repeated displacements in the same direction, saccade execution adapts to land on the expected final position of the target, rather than where it is presented before the saccade (e.g. Deubel, 1995). Rapid saccadic corrections and adaptations demonstrate that the displacement signal is incorporated into visually-guided actions, despite it being inaccessible to conscious awareness. In other words, visual signals that are clearly sufficient to guide motor control and motor learning are nonetheless not sufficient for perceptual experience.

The same dissociation of perceptual experience from attentional and motor control appears to apply to the process of planning, executing, and monitoring saccades and fixations. Although our results suggest we are not able to accurately report on our eye movements, we clearly use information about saccade target selection to guide behaviour over the short and long term. For example, eye movements towards previously-fixated locations tended to be slower and less frequent during extended search, an effect normally referred to as inhibition of return (e.g. Klein & MacInnes, 1999; MacInnes, Hunt, Hilchey, & Klein, 2014). The existence of IOR suggests we store and use a representation of where the eyes have been to facilitate inspection of locations that have not yet been visited. Similarly, in priming of pop-out (Maljkovic & Nakayama, 1994) and contextual cueing (Chun & Jiang, 1998), attention is deployed more rapidly when certain aspects of the search array have been viewed previously, suggesting a stored representation of previous instances of search that can facilitate target detection. Despite the fact that we clearly have, and use, representations of previous instances of visual search, the current study has demonstrated that the ability to explicitly report on this information appears to be minimal, or possibly absent altogether. This is consistent with other instances in the literature of dissociations between our perceptual experience and visually-guided actions in terms of what information they can access and use. Some of these are somewhat controversial, for example, the extent to which visually-guided actions are impervious to visual illusions is the subject of heated debate (e.g. Aglioti, Beltramello, Tassinari, & Berlucchi, 1998; Bruno, 2001; Franz, Gegenfurtner, Bülthoff, & Fahle, 2000; van Zoest & Hunt, 2011). However, it is a relatively well-accepted view that that there is some separation in how and where the visual system stores and represents information that is relevant to action versus that which is relevant for identification, conscious report, and explicit memory (Milner & Goodale, 1995). Klein (1980) suggested that eye movements operate at a level of the visual system that can be separated from attentional processes, inconsistent with premotor theory. Hunt and Kingstone (2003), based on a dissociation in how attention and eye movements were affected by target luminance, concluded that at least some eye movements may be planned and deployed before attentional effects are instantiated. The current results, in demonstrating we have a severely limited ability to introspect on our own eye movements, are similarly inconsistent with attention being deployed to the target of each saccade, presuming that attending to these targets would make them more available for conscious report.

In conclusion, our research has demonstrated awareness of one's own eye movements is extremely limited. We used a converging methods approach, constructing three very different ways to probe awareness. In each experiment, participants clearly used every available strategy to boost their accuracy, but when these strategies were not available, the accuracy with which most participants could identify their own scan path, fixations, or individual eye movement was close to chance. Eye movements play a pivotal role in how we view and perceive our environment by determining our primary visual input, as shaped by the strategies we deploy to search and extract information. These processes rely on information which appears to be accessed by our visual, attentional, and motor systems, but is unavailable to us consciously. This suggests conscious access and control is unnecessary for a wide range of visual, attentional, and motor processes, including perceptual stability across saccades, motor learning, saccade target selection, and inhibition of return. In practical terms, the extent to which people are aware of their own eye movements also has important implications for industry, in which there is an increasing interest in using eye movements in diagnostics, training and interface control. For example, information about eye movements during task performance has been used in aircraft inspection training (Duchowski et al., 2000); interactive graphic display usability (Zhu & Ji, 2004), diagnosis of visual distraction in drivers (Zhang, Smith, & Witt, 2006) and evaluating surgical performance (Tien, Atkins, Zheng, & Swindells, 2010). These approaches rely on the assumption that individuals are consciously aware of their own eye movements, can actively retrieve this information, and can provide accurate reports based on viewing their own scan patterns. The results of the current study seriously undermine this assumption.

Author Contributions

The manuscript was jointly written by all authors. The experiments were jointly designed by ADFC & ARH, and implemented by ADFC. Data were collected by AM & AI,

			Reference	instruction	
		A	B	C	D
	A	0.30 (0.40)	**0.92 (0.41)**	0.24 (0.39)	0.70 (0.37)
Observer	B	**1.60 (0.43)**	0.21 (0.36)	0.33 (0.39)	0.13 (0.35)
Instruction	C	**0.91 (0.43)**	0.33 (0.31)	0.33 (0.34)	0.41 (0.39)
	D	0.66 (0.38)	0.11 (0.36)	0.36 (0.34)	-0.06 (0.36)

Table 3

General linear mixed Effect Model coefficients (and standard error). Comparisons for which the 95% confidence interval does not include 0 are shaded. These can be thought of as significant at an alpha of $p < 0.05$.

and analyzed by ADFC.

Acknowledgements

This work was supported by a James S McDonnell Foundation Scholar Award (ARH). The authors would like to thanks Warren James and Melissa Spilioti for their help with data collection.

Appendix A

Linear mixed-effect model of results from Experiment 1: We ran a general linear mixed-effect model (family=binomial) using the `lm4` package for **R**. Specifically, we fit the model:

$$\texttt{recog} \sim \texttt{reID1 : reID2 + 0 + (1|participant) + (1|image)} \tag{1}$$

We force the intercept to be 0 so that the beta coefficients are more easily interpretable: larger positive coefficients mean that it is easier to successfully recognise your eye movements in the scanpath recognition task, where as negative coefficients mean performance worse than chance. The model fit with standard errors is presented in **Table 3**. The results are in good agreement with the means presented in Figure 3, with the largest effects being seen for the target-landmark (A-B) and landmark-target (B-A) comparisons.

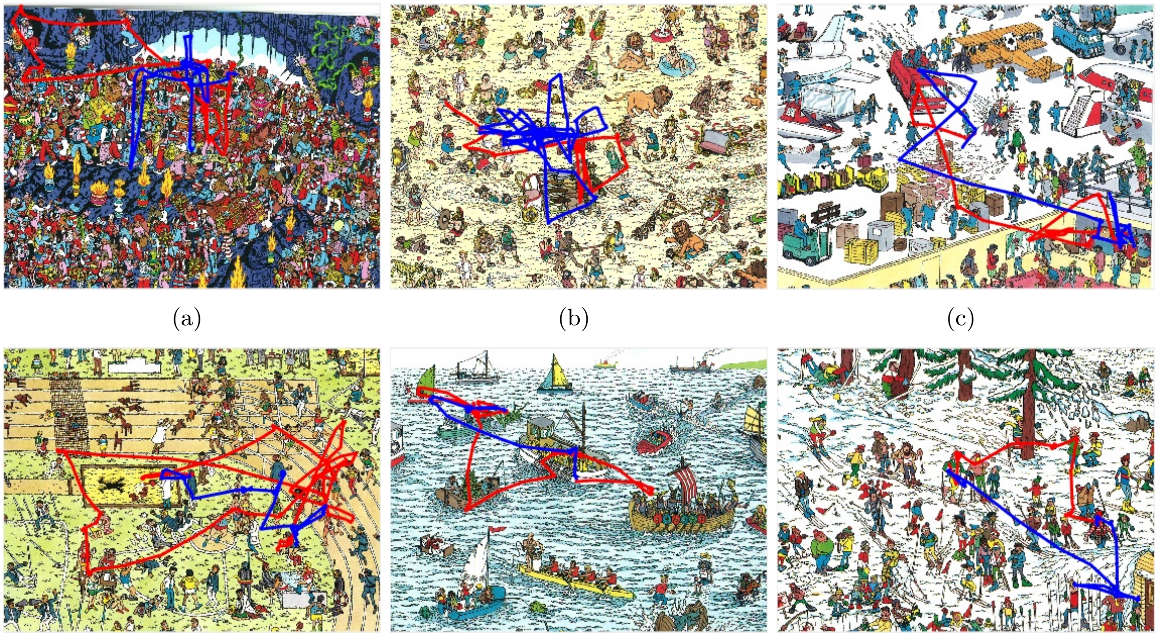
Appendix B

In Experiment 1 we found that participants did slightly better than chance when asked to identify their own scan-path from somebody else searching the same image, but with a different referring expression. When their scan-path was paired with one recorded from another participant following exactly the same referring expression, accuracy was at chance levels. One interpretation of these results is that participants have some limited ability to do an "inverse Yarbus" task, and can occasionally[2] recognise their scan-path by working out which one best fits the task they remember completely.

A potential confound however is that as the structure of the linguistic referring expression varies, the difference between the two scan-paths increases. This raises the question of whether participants' inability to recognise their own eye movements is due to the two

---

[2] even in the easiest condition, accuracy was only at 75%



(a) (b) (c)

(d) (e) (f)

Figure 10. Examples of trials in which the observer's scan-path was paired with one from a second observer carrying out the same condition. In all examples given above, the observer failed to identify which scan-path was their own.

alternative choices being too similar to one another. We explore this possibility in Figure 10 which shows a selection of trials in which the two scan-paths are from observers following the same referring expression, and the participant was unable to identify which scan-path was their own. We can see that even though the two observers were given the same instructions, there are often large differences in where they looked.

References

Aglioti, S., Beltramello, A., Tassinari, G., & Berlucchi, G. (1998). Paradoxically greater interhemispheric transfer deficits in partial than complete callosal agenesis. *Neuropsychologia*, *36*(10), 1015–1024.

Becker, W. & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision research*, *19*(9), 967–983.

Belopolsky, A. V., Kramer, A. F., & Theeuwes, J. (2008). The role of awareness in processing of oculomotor capture: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *20*(12), 2285–2297.

Borji, A. & Itti, L. (2014). Defending yarbus: eye movements reveal observers' task. *Journal of vision*, *14*(3), 29.

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

Bremmer, F., Kubischik, M., Hoffmann, K.-P., & Krekelberg, B. (2009). Neural dynamics of saccadic suppression. *The Journal of neuroscience*, *29*(40), 12374–12383.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision research*, *15*(6), 719–722.

Bruno, N. (2001). When does action resist visual illusions? *Trends in cognitive sciences*, *5*(9), 379–382.

Campbell, F. W. & Wurtz, R. H. (1978). Saccadic omission: why we do not see a grey-out during a saccadic eye movement. *Vision research*, *18*(10), 1297–1303.

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

Clarke, A. D. F., Coco, M. I., & Keller, F. (2013a). The impact of attentional, linguistic, and visual features during object naming. *Frontiers in psychology*, *4*.

Clarke, A. D. F., Elsner, M., & Rohde, H. (2013b). Where's wally: the influence of visual salience on referring expression generation. *Frontiers in psychology*, *4*.

Clarke, A. D. F., Elsner, M., & Rohde, H. (2015). Giving good directions: object ordering reflects visual salience. *Frontiers in psychology.*

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: eye tracking with matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Cumming, G. (2013). The new statistics why and how. *Psychological science*, *25*(1), 7–29.

de Graaf, J. B., Pélisson, D., Prablanc, C., & Goffart, L. (1995). Modifications in end positions of arm movements following short term saccadic adaptation. *Neuroreport*, *6*(13), 1733–1736.

Deubel, H. (1995). Separate adaptive mechanisms for the control of reactive and volitional saccadic eye movements. *Vision research*, *35*(23), 3529–3540.

Duchowski, A. T., Shivashankaraiah, V., Rawls, T., Gramopadhye, A. K., Melloy, B. J., & Kanki, B. (2000). Binocular eye tracking in virtual reality for inspection training. In *Proceedings of the 2000 symposium on eye tracking research & applications* (pp. 89–96). ACM.

Foulsham, T. & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: an experimental test of scanpath theory. *Journal of Experimental Psychology: General*, *142*(1), 41.

Franz, V. H., Gegenfurtner, K. R., Bülthoff, H. H., & Fahle, M. (2000). Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychological Science*, *11*(1), 20–25.

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20–25.

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

Hunt, A. R., Chapman, C. S., & Kingstone, A. (2008). Taking a long look at action and time perception. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 125.

Ibbotson, M. R. & Cloherty, S. L. (2009). Visual perception: saccadic omission-suppression or temporal masking? *Current Biology*, *19*(12), R493–R496.

Johnson, H. & Haggard, P. (2005). Motor awareness without perceptual awareness. *Neuropsychologia*, *43*(2), 227–237.

Johnson, H., van Beers, R. J., & Haggard, P. (2002). Action and awareness in pointing tasks. *Experimental Brain Research*, *146*(4), 451–459.

Klein, R. M. & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological science*, *10*(4), 346–352.

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. Nickerson (Ed.), *Attention and performance* (Chap. 3, pp. 259–276). Hillsdale: Lawrence Erlbaum Associates.

MacInnes, W. J., Hunt, A. R., Hilchey, M. D., & Klein, R. M. (2014). Driving forces in free visual search: an ethology. *Attention, Perception, & Psychophysics*, *76*(2), 280–295.

Maljkovic, V. & Nakayama, K. (1994). Priming of pop-out: i. role of features. *Memory & cognition*, *22*(6), 657–672.

Marti, S., Bayet, L., & Dehaene, S. (2015). Subjective report of eye fixations during serial search. *Consciousness and cognition*, *33*, 1–15.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological bulletin*, *81*(12), 899.

McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics*, *2*(8), 359–362.

Milner, A. D. & Goodale, M. A. (1995). *The visual brain in action*. Oxford University Press.

Mokler, A. & Fischer, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research*, *125*(4), 511–516.

Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2015). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3–25.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1), 31–40.

Robinson, M. M. & Irwin, D. E. (2016). Shifts of attention bias awareness of voluntary and reflexive eye movements. *Experimental brain research*, *234*(6), 1689–1699.

Ross, J., Burr, D., & Morrone, C. (1996). Suppression of the magnocellular pathway during saccades. *Behavioural brain research*, *80*(1), 1–8.

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, *38*(3), 475–491.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: capture of the eyes by new objects. *Psychological Science*, *9*(5), 379–385.

Tien, G., Atkins, M. S., Zheng, B., & Swindells, C. (2010). Measuring situation awareness of surgeons in laparoscopic training. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 149–152). ACM.

van Zoest, W. & Hunt, A. R. (2011). Saccadic eye movements and perceptual judgments reveal a shared visual representation that is increasingly accurate over time. *Vision research*, *51*(1), 111–119.

Watanabe, J., Noritake, A., Maeda, T., Tachi, S., & Nishida, S. (2005). Perisaccadic perception of continuous flickers. *Vision research*, *45*(4), 413–430.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision research*, *48*(20), 2070–2089.

Zhang, H., Smith, M. R., & Witt, G. J. (2006). Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(4), 805–821.

Zhu, Z. & Ji, Q. (2004). Eye and gaze tracking for interactive graphic display. *Machine Vision and Applications*, *15*(3), 139–148.