Re-evaluating Whether Bilateral Eye Movements Influence Memory Retrieval

by

Brady R. Roberts

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2019

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Several recent studies have reported enhanced memory when retrieval is preceded by repetitive horizontal eye movements, relative to vertical or no eye movements. This reported memory boost has been referred to as the *Saccade-Induced Retrieval Enhancement* (SIRE) effect. Across a series of three experiments, memory performance was compared following repetitive horizontal or vertical eye movements, as well as following a control condition of no eye movements. In Experiment 1, we conceptually replicated Christman et al.'s (2003) seminal study, finding a statistically significant, albeit weak, SIRE effect. In Experiment 2, we sought to explore the generality of the effect by presenting to-be-remembered targets auditorily rather than visually during encoding. There was no evidence of a SIRE effect in this experiment: Bayesian statistical analyses demonstrated significant evidence for a null effect. For Experiment 3, we largely returned to the methodology of Experiment 1, except that now horizontal and vertical eye movement conditions were separated into two groups. We again showed no evidence of a SIRE effect, and again there was significant Bayesian evidence in favour of a null. Taken together, these experiments suggest that the SIRE effect as it has been reported in the literature is inconsistent at best or entirely spurious at worst.

Acknowledgements

First and foremost, I would like to thank my supervisors Dr. Myra Fernandes and Dr. Colin MacLeod for all of their feedback, inspiration, and guidance throughout my first years as a budding graduate student. Their unwavering support through a myriad of meetings and manuscript drafts was paramount in the completion of this thesis. This project would also not have been possible without the help of my ambitious research assistants, Tori Dickey, Torin Young, and Carissa Ng. Finally, thank you to Victoria for enduring countless late-night phone calls when I needed to rant about all of the new ideas, frustrations, and excitements that have come with my foray into graduate studies.

Table of Contents

List of Figures

List of Tables

Introduction

Eye movements and memory processes

The field of psychology has long been interested in the role that the human visual system plays in everyday cognition. In 1972, Loftus published an article in *Cognitive Psychology* that reported a positive correlation between the number of eye movements that a participant made while viewing a picture and their subsequent memory for that picture. Since then, researchers have linked eye movements to visuo-spatial working memory (Pearson, Ball, & Smith, 2014), to external projection of mental imagery (Spivey & Geng, 2001), and even direct neuroanatomical connectivity ties to the hippocampus in macaque monkeys (Shen et al., 2016). The study of eye movements is now firmly embedded in the study of cognition, and extensions to the memory literature continue to emerge.

In 2003, Christman and colleagues were the first to investigate a new perspective on eye movements and memory—the potential benefit of bilateral eye movements for memory retrieval (Christman, Garvey, Propper, & Phaneuf, 2003). They undertook this work because of research in their own lab suggesting that bilateral eye movements can enhance interhemispheric interaction (Christman & Garvey, 2000), which in turn was theorized to have a role to play in episodic memory processes (Christman & Propper, 2001; Christman, Propper, & Dion, 2004). Their prediction was that repeated horizontal saccadic eye movements, when performed immediately before a memory test, should cause increased interhemispheric interaction and lead to superior retrieval relative to no eye movements at all. That pattern was in fact what they observed. The first purpose of the current study was to replicate this seminal work by Christman and colleagues, before ultimately moving forward to test the proposed mechanism underlying their observed memorial boost following bilateral eye movements.

Seven years and many articles after Christman et al.'s (2003) work, Lyle and Martin (2010) published the first study that would refer to their finding as the *Saccade-Induced Retrieval Enhancement* (SIRE) effect. Lyle and Martin argued against Christman et al.'s (2003) assertion that the reported memory benefit was due to interhemispheric processes. They employed a letter-matching task both within and across visual fields to provide evidence that the SIRE effect is more likely due to *intrahemispheric* (i.e., within one hemisphere) rather than to interhemispheric (i.e., between hemispheres) activity (Lyle & Martin, 2010). Their idea was that pair detection on the matching task requires intrahemispheric interaction on within-hemisphere trials but mostly interhemispheric processing on across-hemisphere trials. Therefore, the superior performance on within-hemisphere trials that they observed following eye movements suggests increased intra- as opposed to inter- hemispheric interaction (Lyle & Martin, 2010).

Generally, the SIRE effect has been researched using word stimuli and simple memory retrieval tasks (e.g., Christman et al., 2003; Christman et al., 2004). Brunyé and colleagues, however, sought to extend the literature to include picture stimuli (Brunyé, Mahoney, Augustyn, & Taylor, 2009). Further studies have also reported SIRE benefits in autobiographical memory (Parker, Parkin, & Dagnall, 2013), gist-based false alarm recognition (being less likely to falsely recognize critical-lure word associates; Parker & Dagnall, 2007), episodic future thinking (Parker, Parkin, & Dagnall, 2017), and earlier offset of childhood amnesia (Christman, Propper, & Brown, 2006).

While previous research on the SIRE effect has varied widely with regard to its application, the fundamental underlying mechanisms are still in need of clarification. There are currently two major proposed theories that attempt to explain the SIRE effect: Christman et al.'s theory of interhemispheric interaction, and Lyle et al.'s theory of top-down attentional control.

Because the purpose of the current study is to replicate existing work before moving on to further test these leading theories, we will now delve into the background rationale and literature supporting each account.

Interhemispheric interaction

Christman et al. (2003) argued in the very first bilateral eye movement article that their observed memory benefit was due to increased interhemispheric interaction. They based their hypothesis on their own previous work on handedness, which suggested that left-handed individuals have naturally greater interhemispheric interaction due to their larger corpus callosa, and therefore naturally perform better on episodic memory tests (Christman & Propper, 2001; Nielsen et al., 1990). Christman et al. (2003) broadly based their interhemispheric interaction hypothesis of SIRE around the *Hemispheric Encoding and Retrieval Asymmetry* (HERA; Nyberg, Cabeza, & Tulving, 1996; Habib, Nyberg, & Tulving, 2003) and *Cortical Asymmetry of Reflective Activity* (CARA; Nolde, Johnson, & D'Esposito, 1998a) hypotheses. While the HERA model suggests that encoding is mostly left-lateralized and retrieval is right-lateralized, the CARA model alters this slightly to propose that retrieval processes primarily occur either in the right hemisphere or in both hemispheres, depending on the task (Nolde, Johnson, & D'Esposito, 1998a).

Following the logic of the HERA and CARA models, the interhemispheric interaction hypothesis argues that eye movements made before retrieval prime the hemispheres for interhemispheric activity so that tasks taking advantage of bi-hemispheric communication should receive a performance benefit (Christman et al., 2003; Brunyé et al., 2009). According to this hypothesis (and the CARA hypothesis broadly), tasks with more ‘reflective activity’ (i.e., “when retrieval of additional information or more detailed evaluations of activated information are required during episodic remembering”, Nolde et al., 1998a, p. 3513) will lead to the use of both hemispheres at retrieval (Nolde et al., 1998a; Nolde, Johnson, & Raye, 1998b), thus benefiting from SIRE’s interhemispheric interaction priming effect. This proposal has been supported by work from Brunyé et al. (2009), which showed a SIRE benefit when participants performed a YES/NO recognition test, but not when they performed a two-alternative forced choice recognition test (2AFC). Their argument was that the YES/NO test makes use of more reflective activity relative to the 2AFC test, and for that reason showed a performance benefit (Brunyé et al., 2009).

Propper and Christman (2008) also supported the interhemispheric interaction hypothesis, in their case by connecting it to REM sleep processes. It was suggested that because the majority of eye movements made during REM sleep cycles are horizontal (Hansotia et al., 1990), and that there is an increase in interhemispheric coherence during REM sleep (Nielson et al., 1990), the SIRE effect mimics the same processes found in REM sleep, thereby resulting in increased interhemispheric interaction (Propper & Christman, 2008).

Two studies to-date have use neuroimaging techniques to directly assess the interhemispheric interaction hypothesis as an account of the SIRE effect. The first, conducted by Propper and colleagues (2007), used an electroencephalogram (EEG) following bilateral eye movements to show that, relative to controls (who made no eye movements), there was decreased gamma band coherence between the two hemispheres in the lateral eye movement condition. That is, the authors actually observed *decreased* interhemispheric coherence (contrary to the increased interhemispheric interaction that Christman et al., 2003 suggested). They nevertheless argued that *any* changes in neural activity induced by eye movements that resulted in a boost in memory performance still broadly supports Christman and colleagues' (2003) idea that changes in interhemispheric communication is key.

In the second relevant neuroimaging study, Samara and colleagues sought to test the interhemispheric interaction hypothesis using EEG (Samara, Elzinga, Slagter, & Nieuwenhuis, 2011). Contrary to Propper and colleagues' article, Samara et al. found *no* evidence for alterations in interhemispheric interaction following bilateral eye movements across six frequency bands, including gamma (Samara et al., 2011). It should also be noted that Samara et al. (2011) did not find a significant behavioural SIRE effect for neutral valence words although they did find one for emotional words. Ultimately, though, no difference was found in bilateral EEG coherence for either stimulus valence condition. Thus, their results were entirely in contrast to those of Propper et al. (2007) as well as being inconsistent with the broader interhemispheric interaction hypothesis.

Importantly, the interhemispheric interaction hypothesis predicts that only bilateral eye movements (meaning, only horizontal movements) should result in a SIRE benefit. It has been shown that unilateral saccades activate the contralateral hemisphere (Bakan & Svorad, 1969). In the context of SIRE, this would suggest that vertical eye movements should not produce a SIRE effect because they do not produce the same bilateral activation in frontal eye fields as is produced by horizontal eye movements (Habib et al., 2003; Lyle & Edlin, 2015). In summary, the interhemispheric interaction hypothesis predicts that only horizontal eye movements should result in a performance benefit due to the inherent alternating activation of hemispheres. Thus to test this theory, the current study tested measured memory performance following horizontal eye movements compared to both vertical and no eye movements.

Top-down attentional control

In the SIRE literature, a competing account emerged for the observed memorial benefit following bilateral eye movements—that the boost resulted from top-down attentional control processes. Work by Lyle et al. (2008) was the first to suggest that the SIRE effect could be due to the influence of attention. They suggested that bilateral activation of frontal eye fields occurs even with vertical eye movements, such that both types of eye movements should lead to a SIRE effect (Lyle et al., 2008; Rosano et al., 2002; de Haan, Morgan, & Rorden, 2008). Eye movements in any direction lead to frontal eye field and intraparietal sulcus activation, which themselves have been implicated broadly in attention processes (Corbetta & Shulman, 2002; Moore & Fallah, 2004). Because the frontal eye fields are so intertwined with other brain regions within the same attention network (e.g., the central sulcus, cingulate sulcus, medial frontal gyrus, and intraparietal sulcus; Corbetta & Shulman, 2002), the frontal eye fields are theorized to lead to a cascading effect of activation that ultimately results in interhemispheric interaction (Lyle et al., 2008).

As already mentioned, Lyle and Martin (2010) coined the term "SIRE effect." In the same article, they also refined the top-down attentional control hypothesis to further emphasize an attentional component and to eliminate the connection to interhemispheric interaction. Their account suggested that engaging in bilateral eye movements requires top-down attentional processing to successfully perform repetitive purposeful eye movements, thus preparing the participant to make fewer errors, for instance, on subsequent tasks that require higher levels of attentional control (Lyle & Martin, 2010). Further, they proposed that eye movements are tied to activation in frontal eye fields and in the intraparietal sulcus, which are both in turn part of a frontoparietal attention-control network (Corbetta & Shulman, 2002). Since the aforementioned attention network has been implicated in top-down allocation of attention during retrieval, eye movements in any orientation can prime this network. Priming of this network then ultimately aids top-down attention processes—and therefore memory performance—on a retrieval test (Lyle & Martin, 2010).

After finding supporting evidence of the SIRE effect following vertical eye movements (see Lyle et al., 2008), Edlin and Lyle (2013) sought to more directly test their top-down attention control theory. To do this, they employed a typical eye movement paradigm as used by others in the SIRE literature but instead employed a task designed to measure executive attention. Their results indicated that participants who made bilateral eye movements, relative to controls who made no eye movements, experienced greater attentional control as measured by faster reaction times on the *revised attention network test* (ANT-R; Edlin & Lyle, 2013).

Lyle and Edlin (2015) evaluated this attention idea again, opting to employ both horizontal and vertical eye movement conditions, along with a return to more common memory retrieval tests (namely, recall and recognition) instead of the attention test used previously. Critically, some items on the recall test were made more difficult than others by first having participants practise recall from certain studied semantic categories but not others. They then administered a final recall test, with the prediction that items from non-practised categories would be more difficult to retrieve. Lyle and Edlin performed this experimental manipulation because it has been shown in the literature that top-down attentional processes play a key role on memory tests, especially when performance is low and cognitive effort is high (Corbeza, 2008). Results aligned with their predictions: The SIRE benefit was found when memory tasks were harder and therefore required greater top-down attentional control (Lyle & Edlin, 2015). Crucially, they also found a memory performance benefit following vertical eye movements (for a second time, now), which is only predicted by their top-down attentional control hypothesis.

In summary, the top-down attentional control hypothesis suggests that purposeful eye movements (horizontal or vertical) lead to the priming of an executive function attention network that aids the allocation of top-down attentional control at retrieval, ultimately manifesting in a memory performance boost. Much like for the previous theory discussed, the current experiments compared memory performance following each of three critical eye movement conditions (horizontal, vertical, and none) in an attempt to provide evidence in support of or in opposition to the top-down attentional control account.

Handedness Considerations

In addition to eye movement orientation, the SIRE effect has also been reported to be rather sensitive to participant handedness. In fact, Christman and colleagues' (2003) original article based their hypothesis concerning interhemispheric interaction on prior research from their own lab that suggested a similar neural substrate and episodic memory enhancement for participants with positive familial sinistrality (i.e., having left-handed relatives). They argued this was likely due to left-handers' naturally larger corpus callosa, which allows for greater interhemispheric communication (Christman & Propper, 2001; also see Nielsen, Abel, Lorrain, & Montplaisir, 1990). Their idea was that right-handers' naturally smaller corpus callosa could benefit from an increase in interhemispheric communication that comes as a result of bilateral eye movements, whereas left-handed and mixed-handed individuals (with naturally larger corpus callosa) are already at such a high level of baseline interhemispheric communication that the eye movements provide no additional benefit (Brunyé et al., 2009). Because Christman et al. (2003) specifically noted their participants as being strongly-right handed and made the aforementioned case for handedness considerations, most research in this domain has followed suit and similarly only tested right-handed participants (e.g., Brunyé et al., 2009).

SIRE is so sensitive to handedness, in fact, that a brief report by Lyle and colleagues demonstrated—in contrast to the enhancement seen for right-handed individuals—a significant *reduction* in memory performance following bilateral eye movements for left-handed individuals (Lyle, Logan, & Roediger, 2008). This is odd, though, because neither theory of the SIRE effect would have predicted memory attenuation for left-handers.

It is worth noting that Lyle and Edlin's (2015) top-down attentional control theory does not consider right- versus left-handedness; instead, they stated that consistent handedness is the truly important factor. They argued that only consistent handers seem to benefit from the SIRE effect, but unfortunately they did not provide a prediction as to why this handedness inconsistency occurs (Lyle & Edlin, 2015).

In summary, while Christman et al.'s (2003) interhemispheric interaction account predicts that only right-handers should benefit from SIRE, Lyle and Edlin's (2015) top-down attention account makes no differential predictions for right- versus left-handers.

Implications for Eye Movement Desensitization and Reprocessing (EMDR) Therapy

An astute reader may have noticed that the SIRE effect literature employs a paradigm that is very closely tied to the technique used in *Eye Movement Desensitization and Reprocessing* (EMDR) therapy. When EMDR therapy is used to treat *Posttraumatic Stress Disorder* (PTSD)—the most common use for this type of therapy—the patient is typically instructed to make repetitive bilateral eye movements while recalling their traumatic experience (Shapiro, 1989; see Shapiro, 2002, for a review). This is one of the most critical factors warranting further exploration of the SIRE effect: Related therapeutic practises are basing their patient treatment in the efficacy of bilateral eye movements as a memorial tool.

A common problem with implementing therapy techniques for those that suffer from PTSD is that patients' dissociative amnesia of their traumatic memories makes it difficult to recall and work through their trauma in a therapeutic environment (see Samuelson, 2011, for a review). Christman et al. (2003) suggested that the bilateral eye movements made during EMDR therapy play a similar role to the eye movements made in SIRE studies. That is, the role of eye movements in EMDR is to promote eased retrieval of episodic memories, and therefore to allow for therapeutic guidance to take place more effectively (see Propper & Christman, 2008, for a review).

In addition to the reported episodic memory benefits following bilateral eye movements, there is some evidence that these same eye movements can also reduce an individual's emotionality. If this were indeed the case, it would imply that SIRE is even more applicable for PTSD-related therapeutic interventions: Eye movements could allow patients not only to remember their repressed traumatic memories better, but they could also decrease the patient's emotional state, allowing for eased therapeutic guidance to take place. Christman and Propper (2008) reported a study in which horizontal eye movements were associated with a significant neutralization of mood among participants, such that previously happy and previously sad individuals became respectively less happy and less sad following bilateral eye movements. Further, Bartels and colleagues (2018) reported that bilateral eye movements even caused a reduction in sexual fantasy vividness, arousability, and emotionality. In relation to EMDR more specifically, Stickgold (2002) has argued that bilateral eye movements mimic *rapid eye movement* (REM) phase sleep processes. Much like REM sleep, the idea is that purposeful eye movements allow for enhanced cortical integration of memories, thus making the memories less dependent on the hippocampus (and, by extension, on the amygdala), ultimately reducing the emotionality of the memories. Although research on the emotional neutralization effects of SIRE is rather limited, the connection between SIRE and EMDR seems obvious. However, contrasting results and interpretations within both the EMDR (e.g., Greenwald, 1996) and SIRE (e.g., Houben, Otgaar, Roelofs, & Merckelbach, 2019; Lee, Jongh, & Hase, 2019) literatures have resulted in a rather controversial discussion surrounding the efficacy of both effects. Thus, links between the two literatures should be made with caution.

Due to both the theoretical and real-world implications of SIRE for psychological research and therapy applications, the effect has intrigued investigators. First and foremost, therefore, we wanted to better understand the underlying mechanisms that may be at work causing a boost in memory performance. However, it is always wise to begin with replication of a basic phenomena before undertaking further research, so we began with a conceptual replication of the SIRE effect in our laboratory.

Experiments

Experiment 1.

The aim of Experiment 1 was to conceptually replicate the basic SIRE effect in our own laboratory before moving on to investigate the potential mechanisms. Because this was a conceptual replication, we tried to emulate the original bilateral eye movement study by Christman et al. (2003) as closely as possible. However, in contrast to the bulk of the bilateral eye movement literature, we opted to use a within-subject design to increase statistical power; there seemed to be no theoretical basis for the required design choice to alter the effect.

In this first experiment, we asked participants to remember a series of visually presented words, then to undergo a short unrelated filler task, followed by the critical eye movement task, and finally an old/new recognition test.

Method

Materials.

Word lists.

Word lists of concrete nouns were derived from the *International Picture Naming Project* (IPNP) database (Szekely et al., 2004). In each list, average word length was 5.76 characters (*SD* = 1.75) with 1.69 syllables (*SD* = 0.71), and frequency was average (2.85; log-transformed CELEX; *SD* = 1.29). Of the created word lists, one would be studied at encoding for each eye movement condition, while another would serve as lures on the subsequent old/new recognition test. List assignments were fully counterbalanced between eye movement condition blocks, as was their assignment as to-be-remembered or lure stimuli.

Filler task tones

Between study and test, a series of tones was presented as a one minute test-delay filler task, which served the dual purpose of allowing for a later memory test to assess long-term memory, as well as to guard against possible ceiling effects. Participants were asked to press 1, 2, or 3, on the keyboard to identify whether each tone was low (372 Hz), medium (498 Hz), or high (624 Hz), respectively, based on examples provided during the instructions and practice phase.

Eye movement stimuli.

In the horizontal eye movement task, participants were asked to view and visually track a single solid black dot extending approximately 4 degrees of visual angle in diameter, which sequentially appeared in left and right vertically-centered positions on a computer monitor. There was a blank space of approximately 27 degrees of visual angle between the inside edges of the two circle stimuli (see Figure 1).

The vertical eye movement task was very similar except that stimuli now appeared at the top and bottom of the screen along horizontally-centered positions. All visual angle dimensions and stimulus timings were identical to those of the horizontal eye movement task. The centered (no eye movement) task used the same black dot stimulus and timing as in the eye movement tasks, but in this case the stimulus flashed repeatedly at the centre of the screen. A chin rest was used throughout the experiment to ensure consistent visual angles across all participants. Participants' heads rested roughly 46 cm from the screen throughout the experiment.

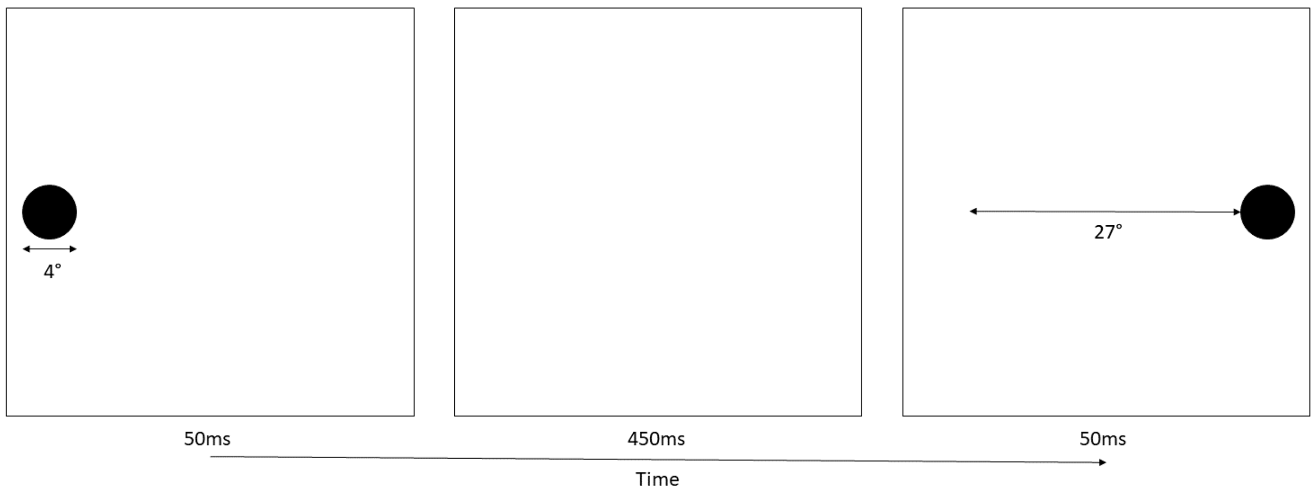

Figure 1. Eye movement phase timing and stimulus size.

Waterloo Handedness Questionnaire.

The Waterloo Handedness Questionnaire (36-item version; Elias, Bryden, & Bulman-Fleming, 1998) was used to determine the extent of participant handedness on a continuum from left to right. Similar to the Edinburgh handedness index commonly used in the eye movement literature (Oldfield, 1971), scores on the Waterloo Handedness Questionnaire can range from -1 (strongly left-handed) to +1 (strongly right-handed), with a score of 0 representing ambidextrous.

Procedure.

Following informed consent, participants completed the *Waterloo Handedness Questionnaire* (WHQ), and were then seated in front of a 24-inch monitor connected to a Windows computer running *E-studio 3.0* (Psychology Software Tools, Pittsburgh, PA). The experimenter then asked the participants to position themselves comfortably in the table-mounted chin rest. The height of the chin rest was adjusted for participant comfort, but was also positioned such that the participant reported being able to stare straight ahead, looking at the centre of the screen. The use of a chin rest in our study is not borrowed from any bilateral eye movement literature, but was employed to ensure consistent visual angle across both horizontal and vertical eye movement trials, for all participants. The experimenter then read all instructions aloud, and participants were offered several opportunities to ask for clarification regarding the experimental procedures both before and after a practice block.

Participants completed one practice block, followed by three counterbalanced experimental blocks. During the practice block, participants completed an encoding phase, a tone-classification filler task, a small sample of each eye movement condition (one after another), and then a short old/new recognition test. Each experimental block was similar to the practise block, with the key differences being longer instances of encoding (2.25 minutes), filler task (one minute), and retrieval phases (one minute), as well as eye movement condition now being blocked (horizontal, vertical, or no eye movements, each for 30 s).

During the encoding phase, 30 words were presented one at a time. A single encoding trial consisted of a fixation cross presented at the centre of the screen for one second, followed by presentation of the to-be-remembered word for three seconds, and finished with a blank screen for 500 ms. Matching the methods outlined in Christman et al. (2003), words were presented in 28 pt upper-case Courier font.

During the post-study maintenance period, the tone-classification filler task was used to guard against potential ceiling effects and to ensure use of long-term memory by minimizing possible recency effect (e.g., Postman & Philips, 1965). During this task, participants were asked to use the keyboard to indicate whether each tone played through the computer speakers was of low, medium, or high pitch. Tones lasted roughly one second and, once a response key was pressed, the next tone played immediately. This continued for one minute.

Following the filler task, participants underwent the three eye movement conditions, one per block. These tasks were designed to replicate previous work by Christman and colleagues (2003), thus stimulus presentation and timing details followed their procedure. In the horizontal eye movement block, participants watched a black dot flash in left and right positions on the monitor for a total of 30 s while following it with their eyes by making repetitive saccades. The vertical block was identical to this except that the black dots now appeared at the bottom and top of the screen. Finally, in the centred (no eye movement) condition, the black dot flashed at the centre of the screen for 30 s and no saccades were made. The black dot stimuli were presented on the screen for 50 ms, with a 450 ms inter-stimulus interval before the next dot appeared, for an overall rate of roughly two presentations (and therefore two saccades) per second.

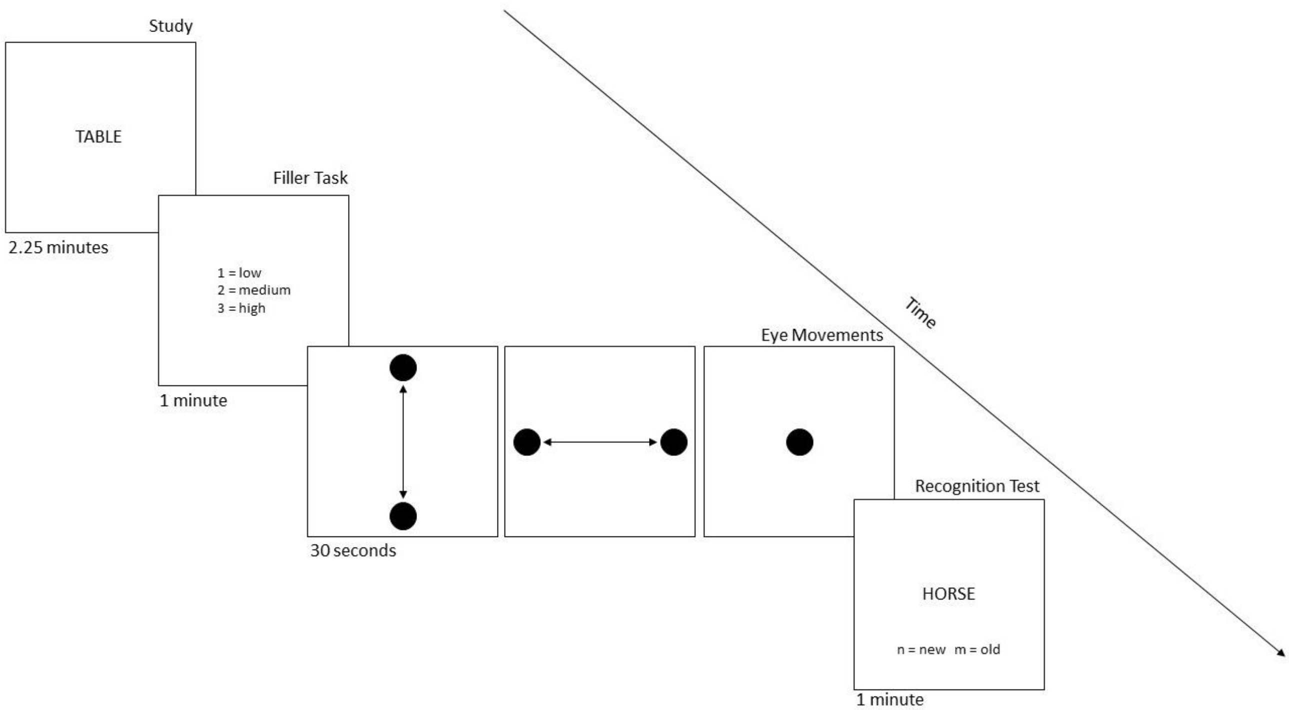

Figure 2. Overall study procedure used for each experimental block. In the critical manipulation, participants experienced one of three eye movement conditions (horizontal, vertical, or no eye movements), as depicted by the black dots with arrows to denote the expected direction of eye movement.

Immediately following the eye movement phase, participants were presented with an old/new recognition test. Thirty words were presented in random order, consisting of 15 randomly chosen target words and 15 randomly chosen lures. Only half of the 30 studied items were presented at test to reduce experiment run time. Participants responded ‘old’ or ‘new’ by pressing keys labeled as such (M or N, respectively on a QWERTY keyboard). Response time and accuracy for each trial were recorded using *E-Prime* Version 3.0 (Psychology Software Tools, Pittsburgh, PA) software. Overall, each recognition test took approximately one minute for participants to complete.

Participants were given a five minute break between experimental blocks to mitigate any possible carry-over effects. During this time, participants were informed that they could leave the chin rest and relax by taking a drink or using their mobile devices. After five minutes, a brief tone was played to notify participants that it was time to get back into the chin rest and resume the study. Once the participants had completed the last experimental block, they were informed of the completion of the study, thanked for their participation, and offered a detailed feedback letter.

All procedures and materials were approved by the Office of Research Ethics at the University of Waterloo (ORE #22799). Data and materials for all of our experiments are available on the *Open Science Framework* at https://osf.io/39uth/.

Participants.

Within the original bilateral eye movement effect article (Christman et al., 2003), an effect size of $d = 0.495$ was found for the critical comparison of horizontal eye movements to no eye movements. We then performed an *a priori* power analysis (matched-pairs, two-tailed t-tests with a $d = 0.495$) using *G*Power* software (Faul, Erdfelder, Lang, & Buchner, 2007), which indicated a required sample size of 35 participants to achieve 80% statistical power. Thus, we aimed to collect 35 participants at minimum in each of our experiments.

In Experiment 1, 42 right-handed University of Waterloo undergraduates (30 female), ranging in age from 17 to 31 ($M = 20.91$, $SD = 2.54$), were recruited to participate for course credit. Participants had all self-reported to be right-hand dominant, to have normal or corrected-to-normal vision, and to have learned English before the age of nine. The Waterloo Handedness Questionnaire (WHQ; Elias et al., 1998) indicated the average handedness score to be 0.65 (SD = 0.17, range = .32 to .99), indicating moderate right-handedness among participants.

Results

Eye movements and memory sensitivity.

Before performing any statistical tests, we analyzed for univariate outliers in the data based on memory sensitivity (d prime; *d'*). By using a critical cut-off of three standard deviations above and below the mean on the memory sensitivity measure, we identified and subsequently excluded data from three participants that were univariate outliers, resulting in a remaining sample size of 39.

Throughout all of the experiments contained within this thesis, d-prime (*d'*)[1] was used as a memory discriminability index, normalized criterion *c* (*c'*) as a measure of response bias, and Cohen's *d* as a measure of effect size. In addition, for consistency, paired-samples Bayesian t-tests were used throughout our analyses. All of these t-tests were two-tailed and used a default Jeffreys-Zellner-Siow (JZS) prior for a medium effect size (Cauchy 0, 0.7071; see Morey & Rouder, 2011; Morey et al., 2015; Rouder et al., 2009). This is the default uninformative prior used in *SPSS* (IBM Corp, Armonk, NY), *JASP* (JASP Team), and the *BayesFactor* (Morey & Rourder, 2015) package for *R* (R Core Team, Vienna, Austria).

Results of the memory test, as measured by memory sensitivity (*d'*), revealed a significant main effect of eye movement condition, $F(2, 66) = 7.28$, $p = .001$, $\eta p^2 = .175$. Bayesian paired-samples t-tests revealed that memory sensitivity in the horizontal eye movement condition ($M = 3.11$, $SD = 0.50$) was significantly greater than in the centered condition ($M = 2.93$, $SD = 0.49$), $t(38) = 2.27$, $SE = 0.08$, $p = .03$, $d = 0.36$, $BF_{10} = 1.40$. The same statistical method also showed that memory sensitivity was significantly higher in the horizontal relative to the vertical ($M = 2.73$, $SD = 0.61$) eye movement condition, $t(38) = 3.92$, $SE = 0.10$, $p < .001$, $d = 0.55$, $BF_{10} = 5.05$, but that the vertical and centered conditions did not differ, $t(38) = 1.88$, $SE = 0.11$, $p = .07$, $d = 0.23$, $BF_{01} = 1.53$. Table 1 displays the means and standard deviations of hits, false alarms, memory sensitivity (*d'*), and response bias (*c'*) on the recognition tests.

[1] Throughout this manuscript, in any case where a participant's number of false alarms equalled zero or number of hits equalled one, we used a standard correction to these values before calculating *d'* scores (Macmillan & Kaplan, 1985). False alarm rates were adjusted to 0.5/N where N is the number of lure trials; similarly, hits were adjusted to (N - 0.5)/N where N is the number of target trials.

Table 1

Experiment 1 means (and standard deviations) for hit rate, false alarm rate, *d'*, and *c'*

Eye Movement Condition	Hit Rate	False Alarm Rate	*d'*	*c'*
Horizontal	.88 (.12)	.08 (.09)	3.11 (0.50)	0.05 (0.12)
Vertical	.77 (.18)	.12 (.13)	2.73 (0.61)	0.13 (0.28)
Centered	.83 (.13)	.12 (.13)	2.93 (0.49)	0.06 (0.15)

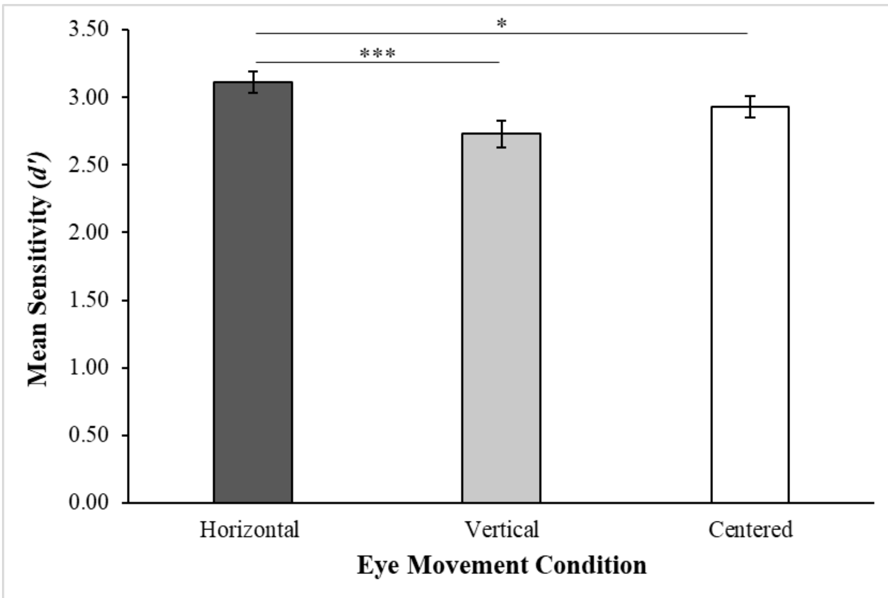

Figure 3. Experiment 1. Mean sensitivity (d prime) on the recognition test following each eye movement condition. Error bars represent ± 1 *SE*. Outlier data are not shown here. * = $p < .05$, *** = $p < .001$.

Discussion

In this experiment, horizontal eye movements led to superior memory relative to both vertical eye movements and a no eye movement control, the latter two of which did not differ from each other. The results of this experiment broadly support the findings reported by Christman et al. (2003). Our findings are also consistent with their interhemispheric interaction account for the SIRE effect—that horizontal eye movements will activate both hemispheres—due to the fact that only horizontal eye movements increased participants' memory sensitivity.

The results are not, however, consistent with the top-down attentional control account of Lyle and Martin (2010). Bayesian evidence for the critical horizontal versus no eye movement comparison was, however, relatively weak ($BF_{10} < 3$; see Jeffreys, 1961). Nevertheless, using classic null hypothesis significance testing standards, we replicated the effect here. To our knowledge, this is the first experiment that has replicated (albeit weakly) the basic SIRE effect using a within-subject design in combination with word-based stimuli.

Experiment 2.

Experiment 1 showed that conceptual replication of the SIRE effect was possible using the word list, test delay time, and within-subject design that was outlined in our Method. Given this, we sought to conceptually replicate again, this time using auditory stimuli. The purpose of this methodological change was to examine the generality of the effect by examining the efficacy of bilateral eye movements in improving memory for auditory stimuli. We also chose to use auditory stimuli because having varied experiment stimuli would allow for increased flexibility in future experiments in which we aimed to investigate the influence of eye movements performed *during* the encoding phase. For instance, we imagined a study design in which eye movements are made in response to visual dot stimuli on the screen (as was the case in Experiment 1), while simultaneously encoding auditorily presented words. Given this plan, the use of auditory stimuli with the SIRE effect required validation before we could address these pertinent questions.

Method

Materials.

The materials used in Experiment 1 were also employed here.

Procedures.

The procedure for the current experiment followed that of Experiment 1, save for three main differences: use of auditory stimuli, removal of the centered eye movement condition, and an (unintentional) reduction in saccade frequency.

During the encoding and retrieval phases, words were presented auditorily one at a time through computer speakers. A single encoding trial consisted of a fixation cross presented at the centre of the screen for one second, followed by a three second period during which a to-be-learned word was presented auditorily. The trial finished with a blank screen for 500 ms. Similarly, at the time of test, words were presented once auditorily and did not advance until a keypress indicating ‘old’ or ‘new’ was made.

The centered eye movement condition was removed to reduce the time and resources needed to run the experiment. We saw the horizontal-vertical comparison as of more theoretical interest, given that it equated factors such as eye movement stimulation and practising overt shifts of attention, leaving only the direction of eye movements as the critical factor driving any subsequent memory benefit.

Finally, saccade rate during the horizontal and vertical eye movement tasks was reduced by roughly half (one saccade per second). This was actually due to an experimenter error during programming. However, examination of the literature produced no theoretical reason that this would alter the effect. Obviously, there would likely be a difference between making one saccade during the task versus one hundred, but the difference here was not nearly as extreme—30 saccades versus 60 saccades.

Participants.

A total of 39 right-handed University of Waterloo undergraduates (24 female), ranging in ages from 18 to 31 (M = 20.54, SD = 2.87), were recruited to participate for course credit. The participant criteria used in Experiment 1 were duplicated here. The average WHQ handedness score of this sample was 0.60 (SD = 0.16, range = -.06 to .94), again indicating moderate right-handedness among participants.

Results

Eye movements and memory sensitivity.

In this experiment, we used the same statistical methodology as in Experiment 1. No statistical outliers were detected in this dataset. Table 2 summarizes the means and standard deviations of the memory scores from the recognition tests.

A Bayesian paired-samples t-test revealed that memory sensitivity in the horizontal eye movement condition did not differ significantly from that in the vertical condition, $t(38) = 0.57$, $SE = 0.14$, $p = .57$, $d = 0.07$, $BF_{01} = 6.83$.[2]

Table 2
Experiment 2 means (and standard deviations) for hit rate, false alarm rate, *d'*, and *c'*

Eye Movement Condition	Hit Rate	False Alarm Rate	*d'*	*c'*
Horizontal	.81 (.14)	.18 (.17)	2.05 (0.98)	-0.16 (-0.02)
Vertical	.82 (.12)	.21 (.19)	1.97 (1.07)	0.29 (0.28)

[2] It should be noted that, prior to this experiment, the authors conducted a very similar study that used the exact same methodology as was used here (reported as a supplementary experiment in Appendix A). Due to experimenter error, however, the study was not counterbalanced properly and therefore a decision was made not to formally report its analyses. Still, it is worth mentioning that a null SIRE effect was found there as well.

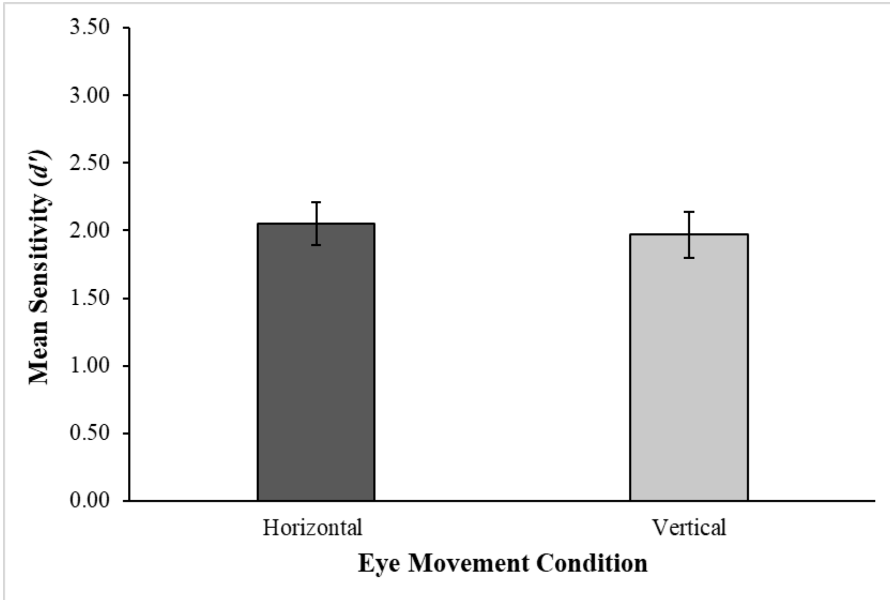

Figure 4. Experiment 2. Mean sensitivity (d prime) on the recognition test following each eye movement condition. Error bars represent ± 1 *SE* .

Discussion

The failure to replicate the SIRE effect in Experiment 2 came as a surprise. Experiment 1 demonstrated that our changes in word lists, test delay, and use of within-subject design were all valid approaches that should result in a significant SIRE effect. Yet here in Experiment 2, Bayesian evidence for a null effect was substantial ($3 < BF_{01} < 10$; see Jeffreys, 1961).

Although Experiment 2 did largely use the same methods as Experiment 1, there are two major differences unique to Experiment 2 that may have led to the observed null effect. First, eye movements made in this experiment were at a rate half that of Experiment 1. This resulted in one eye movement per second, rather than two as was previously used. To our knowledge, no work has been done in the related SIRE literature that would indicate any influence of this factor in determining memory performance. It is the case that most researchers opt to use Christman et al.'s (2003) methodology, which outlines 30 s of eye movements at a rate of two per second. Still, we have not found any research to suggest that number of saccades should matter: It remains an empirical question.

Second, stimuli in the current experiment were presented auditorily at both encoding and retrieval. This presentation format was selected with the intention of replicating the effect using auditory stimuli before using this method in subsequent experiments. Although the present failure to replicate could be due to the use of auditory stimuli, this also does not seem likely. Parker and Dagnall (2007) used a standard bilateral eye movement design with auditorily presented stimuli, and they *did* replicate the SIRE effect for words on a recognition test. Save for different word lists, there are no major methodological differences between Parker and Dagnall's (2007) experiment and our own. Thus, we see no obvious reason for our failure to replicate past research.

Experiment 3.

After our failed conceptual replication in Experiment 2 and weak Bayesian evidence in Experiment 1, we set out once again to replicate the critical SIRE effect. We did not see a benefit in continuing with new experiments using auditory presentation of memory targets until we could assure ourselves that the SIRE effect was indeed replicable in its original form.

To achieve this, we returned to the same methodology as in Experiment 1. The only deviations from Experiment 1's protocol were the number of eye movement conditions experienced by a participant and the length of the recognition test. While in Experiment 1 each participant took part in all three eye movement conditions, in Experiment 3 participants were split into two groups. Group 1 experienced horizontal and centred eye movement conditions, within-subject, while Group 2 experienced vertical and centred eye movement conditions, also within-subject. The purpose of this change was to split up the eye movement orientation types among participants to decrease any effect of interference or carry-over between eye movement conditions. Therefore, this was a mixed design in which the critical comparisons (active eye movement blocks to no eye movement blocks) were made within-subject. The recognition test was also doubled in length to prevent ceiling effects in memory performance.

Method

Materials.

The same materials used in Experiments 1 and 2 were also used here.

Procedure.

Here, we returned largely to the experimental procedures and materials used in Experiment 1, except for the following differences: mixed study design and doubling of recognition test length.

As mentioned above, eye movement conditions were separated in this study for the purpose of guarding against any possible carry-over or interference that might result from switching between eye movement conditions. Therefore, this experiment used a mixed design, with each participant randomly assigned to either Group 1 (horizontal-centred; n = 50) or Group 2 (vertical-centred; n = 51), and eye movement condition being varied within-subject.

In addition, to guard against any potential ceiling effects, the recognition tests were doubled in length and included 30 studied targets as well as 30 randomly chosen lures. Since the recognition test length was doubled, it now took participants approximately two minutes to complete each recognition test.

Participants.

A total of 101 right-handed University of Waterloo undergraduates (79 female) were recruited to participate for course credit; they ranged in age from 17 to 32 (M = 19.31, SD = 2.39). Seven of these original participants' data files had to be excluded due to technical errors and were subsequently replaced with data from new participants. The average WHQ handedness score of this sample was 0.67 (SD = 0.13, range = .25 to 1), again indicating moderate right-handedness. Participant selection criteria were identical to those of Experiments 1 and 2.

Results

Eye movements and memory: sensitivity.

We used the same statistical methodology as in the previous two experiments. Three univariate outliers (± 3 SD) were detected and removed from this dataset, resulting in a final sample size of 49 participants in each of the two groups (total N = 98). Table 3 summarizes the means and standard deviations from the recognition tests.

A Bayesian paired-samples t-test in Group 1 revealed that memory sensitivity was equivalent in the horizontal eye movement condition and the centered condition, $t(48) = 0.13$, $SE = 0.13$, $p = .90$, $d = 0.02$, $BF_{01} = 8.87$. Likewise, in Group 2, memory sensitivity was equivalent in the vertical eye movement condition and the centered condition, $t(48) = 0.21$, $SE = 0.10$, $p = .84$, $d = 0.03$, $BF_{01} = 8.76$.

Table 3

Experiment 3 means (and standard deviations) for hit rate, false alarm rate, *d'*, and *c'*

Group	Eye Movement Condition	Hit Rate	False Alarm Rate	*d'*	*c'*
1	Horizontal	.79 (.18)	.18 (.11)	2.01 (0.97)	0.07 (0.26)
	Centered	.79 (.15)	.18 (.15)	2.01 (1.04)	-0.03 (0.47)
2	Vertical	.75 (.16)	.18 (.14)	1.89 (1.03)	0.11 (0.38)
	Centered	.76 (.17)	.18 (.13)	1.89 (0.94)	0.10 (0.30)

Unique to this experiment, an independent-samples Bayesian t-test was also conducted to compare the horizontal and vertical eye movement conditions between groups. There was no significant difference in memory sensitivity between the two eye movement conditions, $t(96) = 0.71$, $p = .48$, $d = 0.15$, $BF_{01} = 5.09$. Our aforementioned *a piori* power estimates were calculated assuming a paired-samples test, however, so this statistical test was likely not sufficiently powered to detect an effect between-subjects.

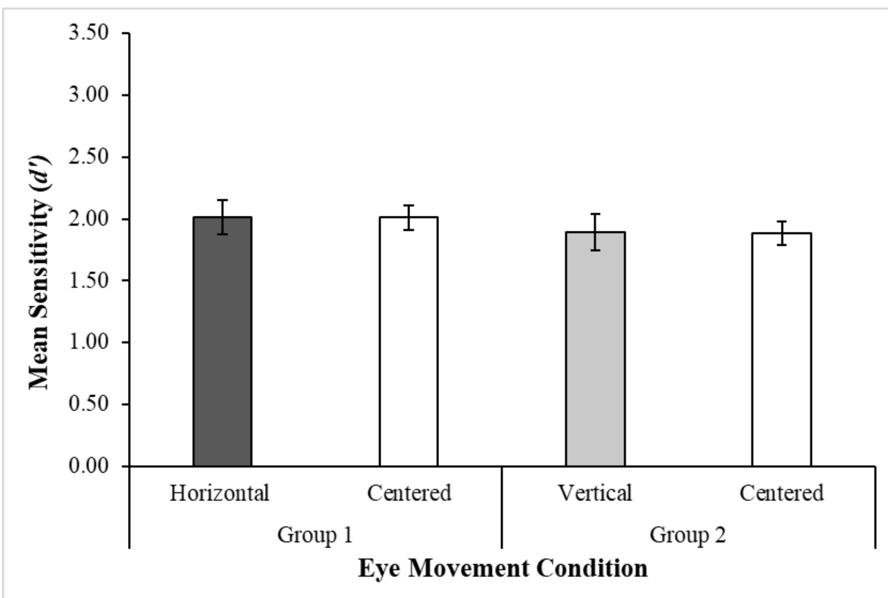

Figure 5. Experiment 3. Mean sensitivity (d prime) on the recognition test following each eye movement condition. Error bars represent ± 1 *SE*. Outlier data are not shown here.

Discussion

Due to weak Bayesian evidence in Experiment 1, and a failure to replicate in Experiment 2, our confidence in the SIRE effect was dwindling. Therefore, in our third and final experiment we sought to return to a more basic and stripped-back experimental design. We employed our largest sample size to-date, combined it with a within-subject methodology to increase statistical power even further (a greater chance of finding a true effect). We also went back to using the literature-based standard of visually-presented word stimuli for the memory task, and a corrected eye movement rate of two saccades per second (the same as in Experiment 1 and Christman et al., 2003). We even went so far as to split up horizontal and vertical eye movement conditions, in case there were any carry-over effects of one saccade orientation on the other.

Despite all these features, Experiment 3 again failed to replicate a basic SIRE effect, with Bayes factors indicating substantial evidence for a null ($3 < BF_{01} < 10$; see Jeffreys, 1961). A failure to replicate here with our increased power, validated methodology, and conceptual replication of past methods, suggests that the SIRE effect as reported in the literature may not exist. That is, the SIRE effect seems to be spurious at best or entirely non-existent at worst. This matter is considered further in the General Discussion.

Participant handedness considerations.

An experienced reader of the SIRE literature may have noticed one difference in our participant samples relative to other work: handedness. It has been argued by some (Lyle et al., 2008, etc.) that participant handedness can have a moderating effect on any subsequent memory boost that bilateral eye movements provide, such that only strongly right-handed individuals should experience a memorial benefit (see the Introduction section for more on this).

Furthermore, Lyle et al. (2008) even showed a small but significant decrease in memory performance following bilateral eye movement in left-handed individuals.

To assess the role of handedness on memory across our studies, we used Bayesian Pearson correlations. Because we used the same measure of handedness across all of our experiments, we were able to collapse for a larger overall sample. Table 4 presents the results of these analyses. As is clear from the table, there was no significant correlation between handedness scores and memory performance within any eye movement condition. In fact, the Bayes Factors demonstrate moderate to strong evidence for a null effect.

Table 4

Bayesian Pearson correlations of Waterloo Handedness Questionnaire (WHQ) scores to memory sensitivity across all 3 experiments.

Eye Movement Condition	Correlation with WHQ Score
Horizontal	$r\,(126) = .06,\ p = .53,\ BF_{01} = 11.77$
Vertical	$r\,(127) = -.11,\ p = .23,\ BF_{01} = 6.91$
Centered	$r\,(138) = -.01,\ p = .93,\ BF_{01} = 14.88$

It is also worth noting that whereas some studies in the SIRE literature demonstrate an effect of handedness on SIRE-related memory performance outcomes (e.g., Lyle et al., 2008), others do not include any metric of handedness information (e.g., Bartels et al., 2018; Propper et al., 2007). An article by Samara et al. (2011) did in fact report handedness of participants but found no evidence for the basic SIRE effect in their strongly right-handed participants (contrary to the prediction of both leading hypotheses), but did find evidence when they used emotionally valenced words. Additionally, Edlin and Lyle (2013) also reported non-significant SIRE by handedness interactions. Finally, Matzke et al. (2015) employed a criterion of only allowing for strongly right-handed participants in their study, and failed to find the SIRE effect. Thus, our work joins others' within the related literature that has found handedness not to be as important a factor to the SIRE effect as some had previously theorized.

General discussion

Overall, the present results do not support the claim that eye movements can lead to improved memory performance. Across three experiments, we found the effects of bilateral eye movements on memory performance to be very small or simply absent. Regardless of the direction of eye movements, there was no consistent change in memory performance observed. Given the appeal of an emergent body of work arguing for the benefit of bilateral eye movements on memory performance, consistent replication is required. Although this project was not a registered report with the goal of directly replicating previous work, the methodology that we employed did closely mimic published work (especially the seminal Christman et al., 2003, article). Importantly, our results do not stand alone; Matzke et al. (2015) also failed to replicate the SIRE effect, finding consistently strong Bayesian evidence for a null. Moreover, their method followed a framework of pre-registered adversarial collaboration, which involved two opposing research groups discussing and ultimately agreeing upon specific experimental and statistical methodologies before pre-registering the study on the *Open Science Framework* (OSF). Still, it is important to clarify any methodological differences and limitations that existed within our own experiments that may have been influential in providing repeated evidence for null SIRE effects.

Methodological considerations

Between- vs. within-subject designs.

One of the largest differences between our work and that of the broader SIRE literature is the type of study design used. While the literature has largely used between-subjects designs; we used a within-subject design. Our design choice was made for the purpose of increased statistical power, given no reason to anticipate that this would alter the effect. Still, when using a within-subject design, one must be wary of carry-over effects. To our knowledge, only two other SIRE articles used within-subject designs. Brunyé et al. (2008) did so and also employed a 10-minute delay between blocks to mitigate carry-over effects. Although they did find a SIRE effect in their study, they did not mention any carry-over effects in their results.

Samara et al. (2011) also used a within-subject design (with a one week delay between sessions) and failed to find the SIRE effect for neutral words (but did find it for emotionally valenced words). Similar to the work by Brunyé and colleagues, counterbalancing was mentioned but no order effects were reported.

Within the present experiments, we used a five minute delay between blocks. We did analyze for order effects but found nothing noteworthy. The broader point both from our experiments and from the related literature remains: Using a between-subjects relative to within-subject design *should not* and *does not* seem to moderate the SIRE effect.

Auditory vs. visual stimuli.

Another departure that we took with regard to the literature was to switch from visual to auditory presentation of words in Experiment 2. After our failure to replicate the SIRE effect in Experiment 2, we thought that perhaps the bilateral stimulation of eye movements needed to occur within the same modality as the stimulus presentation (as was the case in previous studies using bilateral eye movements and visual presentation of words). Upon further investigation, however, we discovered an insightful article by Nieuwenhuis et al. (2013) which used visually presented stimuli but which had a condition in which bilateral stimulation was provided via tactile sensation (alternating tapping of the participants' hands) and a SIRE effect was observed.

Thus, there was at least one case wherein the bilateral activation did not match the modality of stimulus presentation and yet a significant SIRE effect was still reported.

We also discovered an article by Parker and Dagnall (2007), which we believe to be the only other SIRE article that has used auditory word stimuli with a standard bilateral eye movement paradigm. They reported using auditory stimuli played via tape recorder, with two words being presented roughly every six seconds (matching the presentation rate in our Experiment 2). Critically, Parker and Dagnall (2007) reported finding a significant SIRE effect using this method. Taken together, these two studies indicate to us that using auditory relative to visual stimuli should not have undermined observing a SIRE effect for those words.

Looking forward.

It should be clear by now that we intend this article to serve as a cautionary tale with respect to the unreliable SIRE effect, based on our repeated failed attempts at replication. That said, we do *not* intend this as another article preaching the importance of replication in Psychology. We are hopeful that the overwhelming majority of authors in our field are already aware of the ongoing ‘replication crisis’ (e.g., Asendorpf et al., 2013). Thus, it seems more beneficial to discuss potential options for bettering ourselves as a broader scientific community moving forward. More specifically, we believe that researchers within smaller literatures (as in the case of the SIRE effect) should band together to employ the most rigorous replication methods that we have at our disposal.

Matzcke et al.’s (2015) attempted replication serves as a solid foundation upon which to build. Their method of conducting a pre-registered adversarial collaboration is undoubtedly a step in the right direction, but more needs to be done. Ideally, a multisite replication effort could be formed involving leading SIRE effect researchers to provide a high quality, pre-registered replication attempt with high statistical power and agreed upon methods. We believe that this form of collaborative replication would be able to provide heavily weighted evidence in favour of or against the existence of the SIRE effect.

Ultimately, further replication in this literature will only serve to better our field. If the SIRE effect were indeed to consistently replicate—which admittedly on the basis of our work and that of Matzcke et al. (2015) we find an unlikely outcome—we would certainly acknowledge that and retract our criticism. However, even if the SIRE effect fails to replicate consistently, there is still benefit to the scientific community. For instance, potential boundary conditions on the effect can be observed and further studied, perhaps leading to even more interesting theoretical contributions. Often, a failure to replicate can indicate a nuance in methodology or participant demographics that had not been considered before as potentially crucial to observing an effect. If this happens, it has the potential to be empirically interesting and could even inform theory. That is what we hope to promote here—further replications of the SIRE effect in a collaborative manner that benefits everyone.

Concluding remarks

In conclusion, the present thesis describes a broad failure to conceptually replicate the previously reported *Saccade Induced Retrieval Enhancement* (SIRE) effect. Despite some evidence in Experiment 1 that was consistent with existing reports, this evidence was weak. Our next two experiments wholly failed to replicate any SIRE effect. When methodological differences were considered, ultimately these were considered to be inconsequential. Thus, it is evident to us that the SIRE effect, as it has been reported in the literature, is not a consistently reproducible finding. Taken together with other failures to replicate, we must as a rational scientific community question the legitimacy of the effect. Psychology can certainly point to instances of attractive findings that make headlines but then do not replicate. Amidst the current apprehensive climate within our academic community, one should use caution when moving forward with an effect such as SIRE, which arguably fits such a description. We must remember that the work done to date was not in vain; consistent replication attempts can only serve to enhance our empirical base.

References

Asendorpf, J. B., Conner, M., Fruyt, F. D., Houwer, J. D., Denissen, J. J., Fiedler, K., . . . Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*(2), 108-119.

Bakan, P., & Svorad, D. (1969). Resting EEG alpha and asymmetry of reflective lateral eye movements. *Nature*, *223*(5209), 975-976.

Bartels, R. M., Harkins, L., Harrison, S. C., Beard, N., & Beech, A. R. (2018). The effect of bilateral eye-movements versus no eye-movements on sexual fantasies. *Journal of Behavior Therapy and Experimental Psychiatry, 59*, 107-114.

Brunyé, T. T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). Horizontal saccadic eye movements enhance the retrieval of landmark shape and location information. *Brain and Cognition, 70*(3), 279-288.

Cabeza, R. (2008). Role of parietal regions in episodic memory retrieval: The dual attentional processes hypothesis. *Neuropsychologia, 46*(7), 1813-1827.

Christman, S. D., & Propper, R. E. (2001). Superior episodic memory is associated with interhemispheric processing. *Neuropsychology, 15*(4), 607-616.

Christman, S. D., Garvey, K. J., Propper, R. E., & Phaneuf, K. A. (2003). Bilateral eye movements enhance the retrieval of episodic memories. *Neuropsychology, 17*(2), 221-229.

Christman, S., & Garvey, K. (2000). *Episodic versus semantic memory: Eye movements and cortical activation.* Paper presented at the 41st Annual Meeting of the Psychonomic Society, New Orleans, LA.

Christman, S. D., Propper, R. E., & Dion, A. (2004). Increased interhemispheric interaction is associated with decreased false memories in a verbal converging semantic associates paradigm. *Brain and Cognition, 56*(3), 313-319.

Christman, S. D., Propper, R. E., & Brown, T. J. (2006). Increased interhemispheric interaction is associated with earlier offset of childhood amnesia. *Neuropsychology*, *20*(3), 336-345.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215.

de Haan, B., Morgan, P. S., & Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research, 1204*, 102-111.

Edlin, J. M., & Lyle, K. B. (2013). The effect of repetitive saccade execution on the attention network test: Enhancing executive function with a flick of the eyes. *Brain and Cognition, 81*(3), 345-351.

Elias, L. J., Bryden, M. P., & Bulman-Fleming, M. B. (1998). Footedness is a better predictor than is handedness of emotional lateralization. *Neuropsychologia*, *36*(1), 37-43.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Greenwald, R. (1996). The information gap in the EMDR controversy. *Professional Psychology: Research and Practice*, *27*(1), 67-72.

Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: The HERA model revisited. *Trends in Cognitive Sciences, 7*(6), 241-245.

Hansotia, P., Broste, S., So, E., Ruggles, K., Wall, R., & Friske, M. (1990). Eye movement patterns in REM sleep. *Electroencephalography and Clinical Neurophysiology, 76*, 388-399.

Houben, S. T., Otgaar, H., Roelofs, J., & Merckelbach, H. (2019). EMDR and false memories: A response to Lee, de Jongh, and Hase (2019). *Clinical Psychological Science, 7*(3), 405-406.

IBM Corp. (2017). SPSS (Version 25) [Software]. Available from https://www.ibm.com/ca-en/marketplace/spss-statistics.

JASP Team. (2018). JASP (Version 0.10.0) [Software]. Available from https://jasp-stats.org/

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Lee, C. W., Jongh, A. D., & Hase, M. (2019). Lateral eye movements, EMDR, and memory changes: A critical commentary on Houben et al. (2018). *Clinical Psychological Science, 7*(3), 403-404.

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology, 3*(4), 525-551.

Lyle, K. B., Logan, J. M., & Roediger, H. L. (2008). Eye movements enhance memory for individuals who are strongly right-handed and harm it for individuals who are not. *Psychonomic Bulletin & Review, 15*(3), 515-520.

Lyle, K. B., & Martin, J. M. (2010). Bilateral saccades increase intrahemispheric processing but not interhemispheric interaction: Implications for saccade-induced retrieval enhancement. *Brain and Cognition, 73*(2), 128-134.

Lyle, K. B., & Edlin, J. M. (2014). Why does saccade execution increase episodic memory retrieval? A test of the top-down attentional control hypothesis. *Memory, 23*(2), 187-202.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*(1), 185-199.

Matzke, D., Nieuwenhuis, S., Rijn, H. V., Slagter, H. A., Molen, M. W., & Wagenmakers, E. (2015). The effect of horizontal eye movements on free recall: A preregistered adversarial collaboration. *Journal of Experimental Psychology: General, 144*(1), e1-e15.

Moore, T., & Fallah, M. (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. *Journal of Neurophysiology, 91*(1), 152-162.

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods, 16*(4), 406-419.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common design (Version 0.9.12-2) [Software]. Retrieved from https://CRAN.R-project.org/package=BayesFactor

Morey, R. D., Wagenmakers, EJ., & Rouder, J. N. (2016). Calibrated Bayes factors should not be used: A reply to Hoijtink, van Kooten, and Hulsker. *Multivariate Behavioral Research*, *51*(1), 11–19.

Nielsen, T., Abel, A., Lorrain, D., & Montplaisir, J. (1990). Interhemispheric EEG coherence during sleep and wakefulness in left- and right-handed subjects. *Brain and Cognition, 14*(1), 113-125.

Nieuwenhuis, S., Elzinga, B. M., Ras, P. H., Berends, F., Duijs, P., Samara, Z., & Slagter, H. A. (2013). Bilateral saccadic eye movements and tactile stimulation, but not auditory stimulation, enhance memory retrieval. *Brain and Cognition, 81*(1), 52-56.

Nolde, S. F., Johnson, M. K., & D'esposito, M. (1998a). Left prefrontal activation during episodic remembering. *NeuroReport*, *9*(15), 3509-3514.

Nolde, S. F., Johnson, M. K., & Raye, C. L. (1998b). The role of prefrontal cortex during tests of episodic memory. *Trends in Cognitive Sciences*, *2*(10), 399-406.

Nyberg, L., Cabeza, R., & Tulving, E. (1996). PET studies of encoding and retrieval: The HERA model. *Psychonomic Bulletin & Review*, *3*(2), 135-148.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Parker, A., & Dagnall, N. (2007). Effects of bilateral eye movements on gist based false recognition in the DRM paradigm. *Brain and Cognition, 63*(3), 221-225.

Parker, A., Parkin, A., & Dagnall, N. (2013). Effects of saccadic bilateral eye movements on episodic and semantic autobiographical memory fluency. *Frontiers in Human Neuroscience, 7*, 630.

Parker, A., Parkin, A., & Dagnall, N. (2017). Effects of handedness & saccadic bilateral eye movements on the specificity of past autobiographical memory & episodic future thinking. *Brain and Cognition, 114,* 40-51.

Pearson, D. G., Ball, K., & Smith, D. T. (2014). Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition, 132*(3), 416-428.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *The Quarterly Journal of Experimental Psychology*, *17*(2), 132-138.

Propper, R. E., Pierce, J., Geisler, M. W., Christman, S. D., & Bellorado, N. (2007). Effect of bilateral eye movements on frontal interhemispheric gamma EEG coherence. *The Journal of Nervous and Mental Disease, 195*(9), 785-788.

Propper, R. E., & Christman, S. D. (2008). Interhemispheric interaction and saccadic horizontal eye movements: Implications for episodic memory, EMDR, and PTSD. *Journal of EMDR Practice and Research*, *2*(4), 269-281.

Psychology Software Tools, Inc. (2016). E-Prime (Version 3) [Software]. Available from https://www.pstnet.com.

R Core Team. (2013). R: A language and environment for statistical computing (Version 3) [Software]. Available from http://www.R-project.org/.

Rosano, C., Krisky, C. M., Welling, J. S., Eddy, W. F., Luna, B., Thulborn, K. R., & Sweeney, J. A. (2002). Pursuit and saccadic eye movement subregions in human frontal eye field: A high-resolution fMRI investigation. *Cerebral Cortex, 12*(2), 107-115.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237.

Samara, Z., Elzinga, B. M., Slagter, H. A., & Nieuwenhuis, S. (2011). Do horizontal saccadic eye movements increase interhemispheric coherence? Investigation of a hypothesized neural mechanism underlying EMDR. *Frontiers in Psychiatry, 2,* 4.

Samuelson, K. W. (2011). Post-traumatic stress disorder and declarative memory functioning: A review. *Dialogues in Clinical Neuroscience, 13*(3), 346-351.

Shapiro, F. (1989). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress*, *2*(2), 199-223.

Shapiro, F. (2001). EMDR 12 years after its introduction: Past and future research. *Journal of Clinical Psychology, 58*(1), 1-22.

Shen, K., Bezgin, G., Selvam, R., Mcintosh, A. R., & Ryan, J. D. (2016). An anatomical interface between memory and oculomotor systems. *Journal of Cognitive Neuroscience, 28*(11), 1772-1783.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*(4), 235-241.

Stickgold, R. (2001). EMDR: A putative neurobiological mechanism of action. *Journal of Clinical Psychology, 58*(1), 61-75.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250.

Witelson, S. (1985). The brain connection: The corpus callosum is larger in left-handers. *Science, 229*(4714), 665-668.

Appendix

Appendix A: Supplementary experiment statistical analyses

Participants.

A total of 29 right-handed University of Waterloo undergraduates (21 female) were recruited to participate for course credit; they ranged in ages from 18 to 24 (*M* = 20.32, *SD* = 1.87). The average WHQ handedness score of this sample was 0.71 (*SD* = 0.16), indicating moderate right-handedness. Participant selection criteria were identical to those of Experiments 1, 2, and 3.

Results.

Eye movements and memory: sensitivity.

We used the same statistical methodology as in the three reported experiments. Two univariate outliers (± 3 *SD*) were detected and removed from this dataset, resulting in a final sample size of 27 participants.

An analysis of memory sensitivity (*d'*) revealed a non-significant main effect of eye movement condition, $F(2, 52) = 0.15$, $p = .861$, $\eta p^2 = .006$. Critically, a Bayesian paired-samples t-test revealed that memory sensitivity in the horizontal eye movement condition (*M* = 3.01, *SD* = 0.45) was not significantly different than in the centered condition (*M* = 3.00, *SD* = 0.64), $t(26) = 0.11$, $SE = 0.11$, $p = .92$, $d = 0.02$, $BF_{01} = 6.70$. Additional Bayesian paired-samples t-tests revealed that memory sensitivity did not differ significantly in the horizontal relative to the vertical (*M* = 3.05, *SD* = 0.52) eye movement conditions, $t(26) = 0.45$, $SE = 0.09$, $p = .66$, $d = 0.08$, $BF_{01} = 6.11$, nor in the vertical relative to the centered conditions, $t(26) = 0.53$, $SE = 0.10$,

p = .60, d = 0.08, BF_{01} = 5.87. Table 5 displays the means and standard deviations of hits, false alarms, memory sensitivity (*d'*), and response bias (*c'*) on the recognition tests.

Table 5
Supplementary experiment. Means (and standard deviations) for hit rate, false alarm rate, d', and c'.

Eye Movement Condition	Hit Rate	False Alarm Rate	d'	c'
Horizontal	.86 (.11)	.10 (.10)	3.01 (0.45)	0.05 (0.14)
Vertical	.85 (.13)	.14 (.16)	3.05 (0.52)	-0.03 (0.28)
Centered	.81 (.15)	.19 (.19)	3.00 (0.64)	0.08 (0.28)

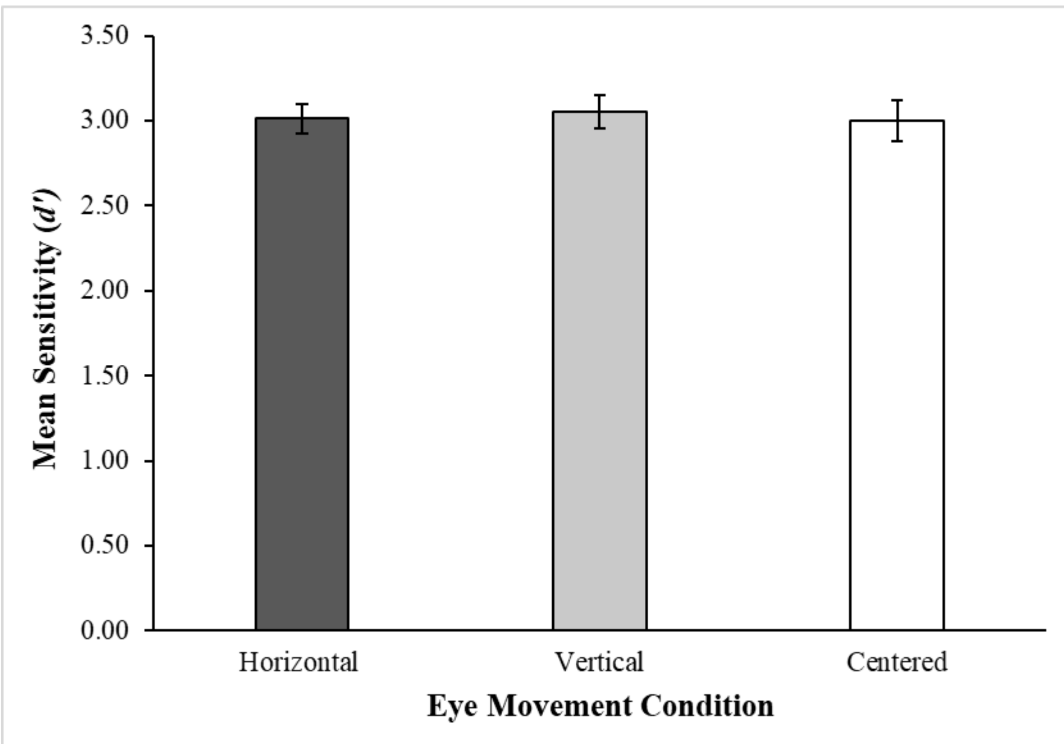

Figure 6. Supplementary experiment. Mean sensitivity (d prime) on the recognition test following each eye movement condition. Error bars represent ± 1 *SE*. Outlier data is not shown here.