

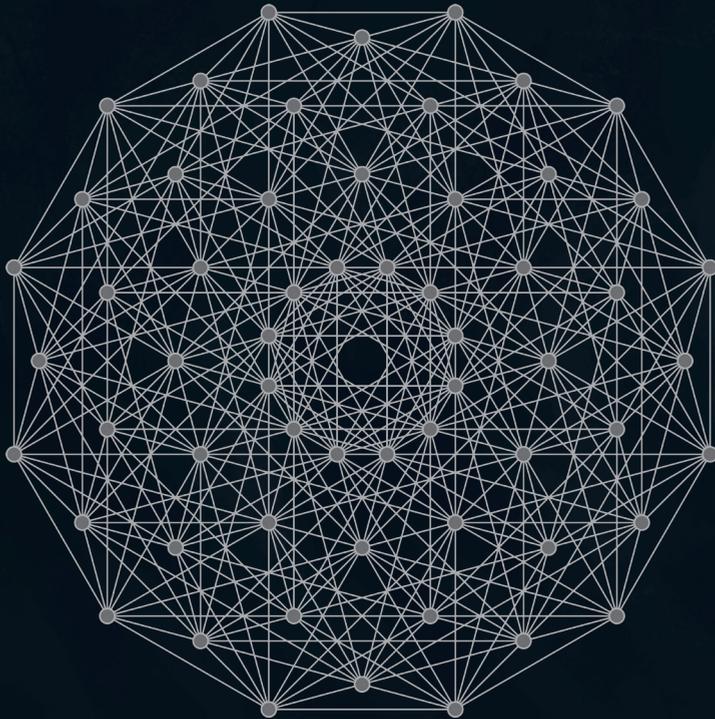


Paul Zerr

THERE AND BACK AGAIN

Memory, eye movements,
and psychophysics on three continents

PhD Thesis



There and Back Again

Memory, eye movements, and psychophysics
on three continents

Darheen en Weer Terug

Geheugen, oogbewegingen en psychofysica
op drie continenten

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
woensdag 9 februari 2022 des middags te 12.15 uur

door

Paul Zerr

geboren op 20 mei 1987
te Frunse

Promotor:

Prof. dr. S. van der Stigchel

Copromotor:

Dr. S. Gayet

The studies in this dissertation were funded by the Netherlands Organisation for Scientific Research (NWO, The Hague, grant number 45213008 to S. van der Stigchel and 191G.085 to S. Gayet).

The artwork on the front cover (Jorm S), lower back cover (twentysecond), on pages 11 (twentysecond), 47 and 193 (Olchik Illustrates) was licensed to the author via shutterstock.com.

The artwork on the upper back cover is public domain and was downloaded from https://commons.wikimedia.org/wiki/File:Rectified_hexacross.svg.

The dissertation was printed by proefschriftmaken.nl.

Typesetting and design was performed by the author.

Cover design was performed by Natalia Vish and the author.

“We have to remember that what we observe is not nature herself, but nature exposed to our method of questioning.”

Werner Heisenberg

TABLE OF CONTENTS

FOREWORD.....	9
CHAPTER 1: INTRODUCTION	11
1.1 Prologue	13
1.2 Vectors, dark rooms, and double step saccades	17
1.3 Remapping fragile memory.....	21
1.4 120 hours of oriented rectangles	29
1.5 Eye movements in the formally blind	33
1.6 Serial dependence	39
1.7 Epilogue.....	43
1.8 References	45
CHAPTER 2: RESEARCH ARTICLES.....	47
2.1 Error compensation in random vector double step saccades with and without global adaptation.....	49
2.2 Remapping high-capacity, pre-attentive, fragile sensory memory.....	75
2.3 The development of retro-cue benefits with extensive practice: implications for capacity estimation and attentional states in visual working memory	97
2.4 Successful visually guided eye movements following sight restoration after congenital cataracts	125
2.5 Serial dependence reflects integration of task-relevant information in working memory	163
CHAPTER 3: DISCUSSION.....	193
3.1 What have we learned? A brief summary.....	195
3.2 Samenvating in het Nederlands	199
3.3 The perpetual diaspora of early career researchers.....	203
3.4 Fundamental vs. applied science	207
3.5 Why then, would a rational agent choose this path?	209
ACKNOWLEDGEMENTS.....	211

FOREWORD

Six years and one apocalypse later, I finally hold in my hands the written record of a life-defining journey. On these pages I describe my progression through the various research projects completed during the course of my PhD. Due to the generous nature of my supervisors and my own drive to learn new skills and theoretical frameworks, my work has been rather diverse, each publication a thesis in itself. Thus, the scientific deliberations are left to the individual research articles, as these manuscripts themselves comprise the necessary considerations. Each research article has its own introduction and discussion that pertains to the individual projects.

I was fortunate enough to choose my own research topics, as a result of which my progress was self-determined, and I could work according to my interests. Each project represents not only a scientific publication, but also an aggregation of experiential learning, fuelled by the passion of curiosity. Each project involved dealing with new data structures, experimental paradigms, statistical methods and different theoretical, programmatical and mathematical challenges. These building blocks together now form the house of my scientist persona.

In this thesis I will first lay out the crucial elements of my journey. I will describe what led me to embark on each project's trajectory, how I decided to research that topic in particular and not any other. As a doctorate is both a professional occupation and a process of learning (indeed, science by definition is a process of perpetual learning), I will also describe what each project taught me.

The introduction is followed by re-prints of each manuscript produced during the PhD, four of which are already published, and a fifth, which is currently under review. I will conclude the thesis with a brief summary of my findings and discuss some of the thoughts that arose during these years pertaining to life in academia.

CHAPTER 1:
INTRODUCTION
There and back again



1.1 Prologue

One questionable morning the servers of Groningen University crashed under a whole generation of Bachelor students desperately and simultaneously signing up for their thesis projects. I found a place with Prof. Jan Visser, who offered a project in the realm of cognitive placebo effects. I ended up applying sham transcranial current stimulation (tDCS) to oblivious participants playing a computer game while their heart rate was being measured. Would believing to undergo a tDCS intervention be sufficient to increase cognitive performance, or even impact a questionable measure of 'cognitive effort' in the heart rate variability data? Of course, as Bachelor theses go, the measurements were poorly calibrated, the theory was convoluted, rested on layers of unproven assumptions, and was mostly held together with faith and duct tape. This accomplished with full success the primary objective of the task: to take my first naïve steps into research territory and to make some low-cost mistakes. No significant effects were found, and I learned a great deal about running and analysing experiments, physiological measurements, current stimulation, the placebo effect, cognitive performance, and mental effort. Most importantly, I got a glimpse of what it can be like to do research with a kind and helpful supervisor.

It probably was then that I decided (despite warnings) to pursue the path of academic research and applied to the Neuroscience & Cognition Research Master at Utrecht University. I was accepted to two other research master programs but went with Utrecht because of their focus on research internships. I had enough of the type of schooling where a lecturer would be reciting the textbook with a focus on facts rather than understanding. Utrecht offered something else. An intense and diverse 10-week course in the fundamentals of neuroscience, followed by a 9-months and a 6-months research internship. This was exactly what I wanted: to do real, hands-on research.

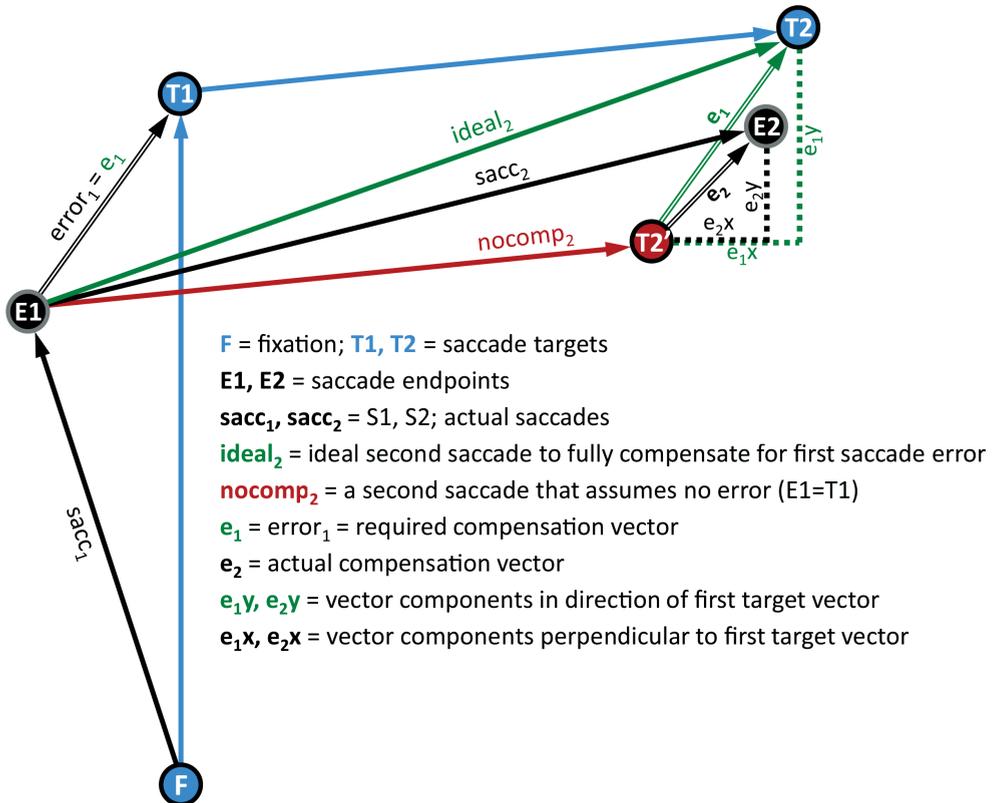
One researcher told me when I met him about eye movements and eye trackers, how much of the visual system deals in gaze centred coordinates, and how the location of any object maintained in memory has to be recomputed with every eye movement (spatial remapping or

updating). I never was overly interested in eye movements, nor did I care about the process of spatial remapping. What I really wanted to do was dive into electrophysiology, be that in slices of mouse brains or human EEG data. What immediately caught my attention, however, was the researcher himself, who appeared friendly, helpful, enthusiastic, and knowledgeable. From the first moment we just clicked and since at the time I didn't know anyone in Utrecht who was focused on brain oscillations, I decided that I might as well try something entirely different.

Little did I know that this researcher would eventually hire me for the long journey that became my PhD. Before this came to pass, however, I did my second internship at the Donders Institute in Nijmegen, where I followed my initial dream and applied theta wave transcranial current stimulation (tACS) on the frontal-midline region of healthy human participants to entrain and boost brain oscillations hypothesized to modulate risk and reward behaviour in people using a gambling task. The project was already set up and was responsible for collecting the data. This meant that I applied tACS and EEG electrodes on people's scalps, have them do the task and later analyse EEG and behavioural data. I spent many, many hours on this task, manually classifying artefacts and aggregating data. Meanwhile I was also fortunate enough to be able to pilot on myself a protocol I wrote to measure EEG while near-simultaneously applying tACS pulses to my own brain. The concept was to alternate between 60 seconds of stimulation and 60 seconds of EEG measurement. I was hoping that this interleaved protocol would circumvent the problem of the stimulation current interfering with the measurement of scalp-derived brain waves, as the stimulation was several orders of magnitude higher than the few μV that could be picked up in the EEG. I did find a small increase in alpha power, and while this project never went anywhere much, it was great fun to do and together with the main internship project satisfied that urge in me to deal with brain electrophysiology. Finally, I was also taking a brain computer interface course led by Dr. Jason Farquhar. We learned about brain computer interfaces in a highly applied and engaging setting. As final project, I programmed and successfully tested a neurofeedback application, in which I and two other volunteers managed to obtain control over our occipital alpha oscillations after about ten hours of practice. I am immensely

grateful for my time in Utrecht, Nijmegen and Hyderabad, and the many learning experiences I was exposed to.

1.2 Vectors, dark rooms, and double step saccades



Towards the end of the research master there was an open call for a PhD position at the Helmholtz Institute, Experimental Psychology, Utrecht University, where I did my first internship. I answered and after some rounds it came down to two candidates, the other being a brilliant fellow N&C student. Our future supervisors and other professors at the department believed in our potential too much to make such a decision and instead were able to pool funds to make two positions out of one in order to hire both of us. I still remember the moment when I received, after months of waiting, the final 'yes'. I was

in shock, disbelief, extasy and awash in excitement. In September 2015 I began working in the Langeveld building at Utrecht University as PhD student of (now Prof.) Stefan Van der Stigchel.

The first project was a continuation of my internship work, in which I investigated how location information is represented in working memory before and after an eye movement. Much of the visual system encodes location in retinotopic coordinates, that is, relative to the centre of gaze. In such a system, coordinates must be recoded (remapped) with every eye movement because of the change in reference point. This can easily be appreciated by moving the eyes from left to right. All objects in the visual field have moved, yet we do not experience them as moving, to us they are still in the same place in the world. Considering that humans execute several eye movements per second, this remapping process must be highly efficient in order to keep up with the processing demand, and has been an area of study for a long time. The investigation of the mechanisms involved in spatial remapping often makes use of the “double step paradigm” (McLaughlin, 1967). In this task, two locations on a screen are briefly shown and after they disappear, participants are required to make two eye movements, to the first and then the second location previously shown. When performed in a dark room with nothing else visible, this demands a tricky computational operation. The first eye movement can simply be executed towards the first remembered location. For the second eye movement, however, the brain needs to know two things: where the eyes are currently pointing, and where the second target was located relative to current eye position. Both can be inferred from the first eye movement. If the brain knows the executed eye movement vector exactly, it can calculate how far the eyes have rotated and thus what the vector between current eye position and second target should be. There is only one problem: biological movements are rarely entirely exact. If the first eye movement is off by any distance, a pre-computed second eye movement vector would miss the second target by that amount. This was my first research question: are landing point errors during the first eye movement compensated for during the second eye movement even in the absence of any visual reference?

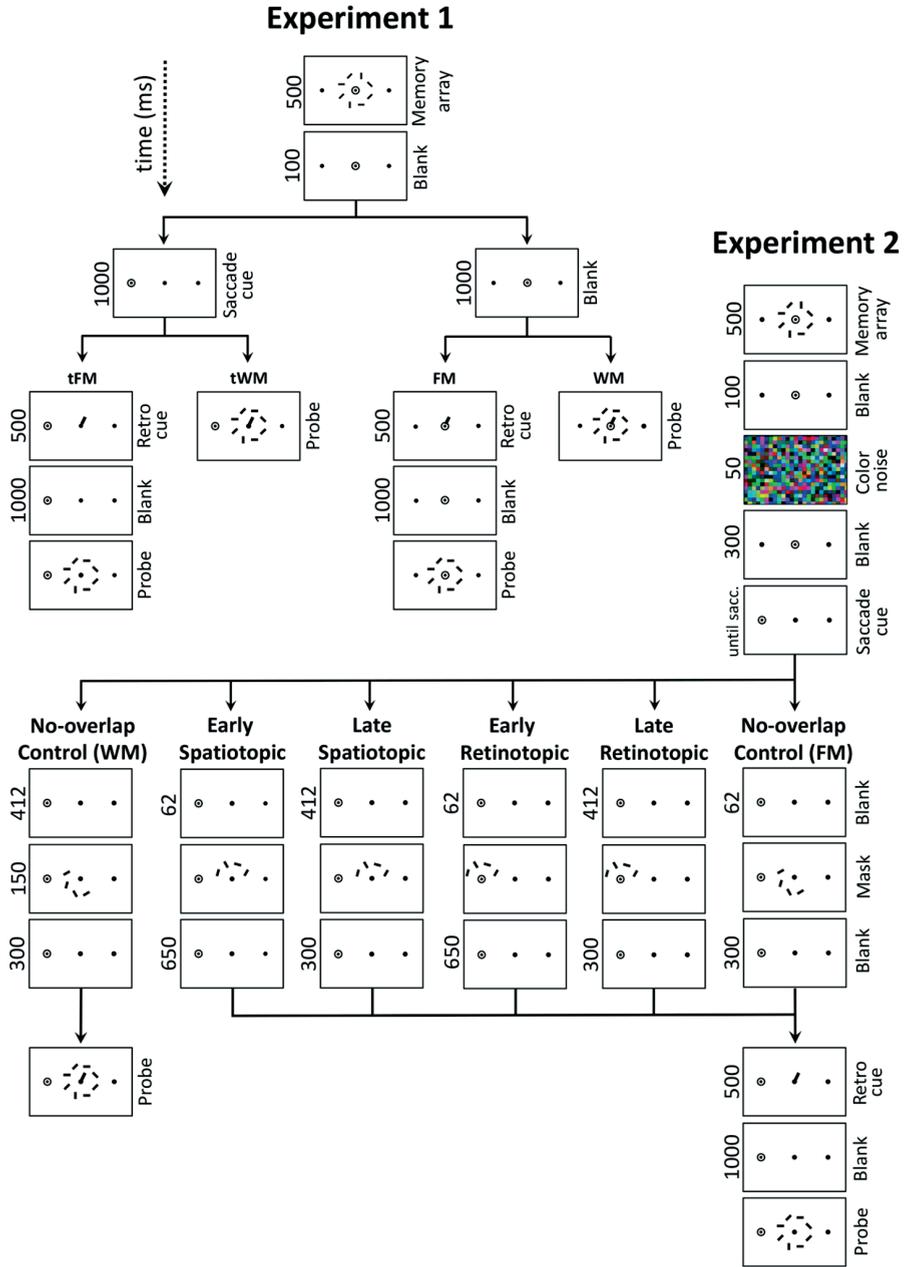
My first task was to create a condition of absolute darkness. Thankfully the lab rooms were endowed with solid doors and no light would enter from the outside. However, a room full of testing

equipment, computers and eye trackers necessarily contained numerous LED's and other light sources, which had to be carefully taped over. Next, a screen had to be manifested which would allow me to display two dots of light and nothing else. I found an ancient Phillips CRT monitor on ebay with physical knobs for brightness and contrast control, and in combination with some neutral density foil I was able to remove both the residual background illumination and afterglow of the monitor. An interesting and difficult challenge! In the end I succeeded in creating what I wanted: a condition of perfect darkness with tightly controlled tiny spots of light displayed on a screen. With some help from Edwin Dalmaijer, I proceeded to program my first experiment in Python. This involved implementing gaze-contingent displays, as in a double step paradigm the second dot had to be removed from the screen at the exact moment the eye movement towards the first dot started. This required efficient programming as the eye movement position had to be read from the eye tracker and translated into a display change before the eyes had reached their first target. Otherwise, the mismatch between predicted and actual retinal position of the first target could be used as reference to program the second eye movement. I needed participants to fully rely on an internal tracking of eye movement commands (known as efference copy or corollary discharge).

As I was piloting the experiment and finally running it on consenting participants, I started to get a feel for the practice of eye tracking and received my first real, useful data. I soon realized that I had set myself up for a different challenge. While reading out the position information from the eye tracker output was comparatively easy, analysing the data in the context of the experiment was not. One of my reasons for running this particular experiment was what I had observed in the remapping literature using the double step paradigm: in these experiments, participants were asked to make extremely repetitive eye movements. In the older literature this was so stereotyped that participants always made the same movements: e.g., from the centre to the right, and then upwards. Back to centre and repeat. More recent investigations allowed for a little bit more variation: while again always starting at the centre, participant's eyes moved up, down, right or left, and then to one side at right angles (see also section 2.1, Figure 2). I was not satisfied with the highly predictable nature of these eye movement patterns. What if results

could be explained by an overlearned, stereotyped notion of where to move the eyes after having done so a few hundred times? I wanted to be able to show targets unpredictably, on any location of the screen. While this was no problem in the implementation of the experiment, it proved to be somewhat more involved during data analysis. Linear algebra was not at all fresh in my mind, having been out of school for many years. So, I started to learn again about vectors and trigonometry. I began to understand why the literature chose the simpler right-angled display option. What I also realized, was that while in school these topics were covered with the notorious grey cloth of classroom education, I now had a goal in mind, a purpose to pour myself into, a problem to solve. This imbibed me with a hunger to understand and an endless energy to bite my teeth into the subject matter until my neural networks wrapped themselves around a semblance of applicable understanding. With the generous help of Siarhei Uzunbajakau, who was working as a lab technician at the department, I was eventually able to find a way of analysing my data appropriately, despite the more complicated random stimulus displays. Not only did I refresh my mathematics, but this kindled in me a great joy for difficult problems that were just outside my current expertise without being too overwhelming to be impossible to solve. I am very grateful to Siarhei, who sat with me for hours, out of no other consideration but the desire to help and a shared love for understanding. As lab technician he was under no obligation to spend time on this with me but did anyway. In the end I insisted on making him a co-author on the paper for his invaluable contributions. This process showed me another aspect of academic science that I remember to this day: that we are all here to learn, to understand, and to help each other in doing so. There are many more scientific elements to this project, which I will not describe here but refer the reader to the re-print of the paper in section 2.1, which was eventually accepted for publication in Vision Research.

1.3 Remapping fragile memory



In the previous chapter I described my investigation into spatial remapping from a low-level, technical perspective, in a task requiring only one or two locations to be kept in memory at a time. Looking up from my computer screen and out the window into an assortment of trees, it is immediately obvious that perception is not limited to just one or two objects, isolated in perfect darkness. Perception appears very detailed from a subjective perspective. While subjective experience is unlikely to be a very reliable estimator of an underlying brain process, perception as we experience it is undeniably real in itself. While there are many different ways in which perception can be implemented (none of which may be intuitively understandable, such as for example rapidly switching attentional spotlights), what I see necessarily is really what I see, and no cognitive scientist can take that away from me. With this in mind, it is striking that lab experiments entirely fail to capture this richness of subjective experience.

In the words of Lamme (2010): “Introspectively, consciousness seems rich in content [...] From the third-person perspective of the behavioural scientist, however, consciousness is rather miserable”. Cohen et al. (2016) write: “Why do we think we see so much when the scientific evidence suggests we see so little?”. They answer this by saying “We consciously perceive more than we can attend, remember, report, or base decisions on. There is no tension between our subjective impression of the world and objective measures of human capacity limits because both of these are true.”. In other words: lab experiments typically require participants to press a button in response to perceived stimuli. This requires a long chain of events to take place between the physical event of the stimulus and a resulting behaviour: from reception, processing, perception, decision making to motor planning and execution. Often it is even required to maintain a representation of the stimulus in memory for some time. In this type of experiment, it is only possible to measure that which survived all these steps and made it into a recorded measurement in some way. Whether that data point takes the form of a behavioural memory reproduction response, perceptual decision making response, a response bias or even neuroimaging (which at this point in time is still noisy and low-resolution), the experimenter has to rely on an indirect measurement of a cognitive process. We therefore have to be very careful when generalizing scientific understanding to the actual human experience. We are not studying human cognition, but

correlates of a construct of cognition as it is taking place in the laboratory. That being said, I strongly believe in the value of the endeavour of cognitive scientists to understand more about the biological and computational machinery that enables us to do things such as perceiving the marvellous beauty of the natural world, effectively navigating a car through traffic, as well as writing and reading these very sentences. Everything we do and experience relies on our brain. Understanding the human brain and mind is as crucial to the continued survival of our species as understanding laws of the external physical world.

I was fascinated by a paper published by Dr. Ilja Sligte (Sligte et al., 2008), a researcher at the Vrije Universiteit Amsterdam. He had discovered, that while traditional paradigms of measuring working memory consistently produce capacity estimates of around three to four objects, a small tweak in the paradigm upped that estimate dramatically to 16 objects or more. This tweak had been known, in a related form, since at least 1960 when Sperling published a seminal paper on something called iconic memory, using the partial report paradigm. In this task, participants would briefly see a large array of objects, which they were asked to remember. Very shortly after the offset of the array a cue would indicate a subset of the array. An object from this subset would later have to be retrieved from memory. This means, that in order to complete this task correctly, participants had to briefly remember all objects in the array (until seeing the cue). After seeing the cue they could focus their memory resources on a small subset of objects and discard the other objects from memory. When using this partial-report cue, participants could report a great many objects, which clearly demonstrated that they could remember and access more than four objects, at least for a very brief period of time. This is necessarily the case because otherwise the cue would not allow them to selectively continue to remember a subset. When the cue was shown later than a few hundred milliseconds after the memory targets disappeared, no such benefit was obtained versus not showing any cue. This led to the description of “iconic memory” (Sperling, 1960), a high-capacity and very short-lived form of short-term memory. This was contrasted with “working memory” (e.g., Baddeley & Hitch, 1974; Averbach & Coriell, 1961), a low-capacity, durable and highly persistent form of short-term memory. Over time, however, iconic

memory fell out of favour, and was largely seen as an artefact of the system, rather than a functional element of cognition.

In the early 2000s, two labs revived the concept of partial-report, which would then become known as the “retro-cue paradigm” (Landman et al., 2003; Griffin and Nobre, 2003). Contrary to iconic memory, in this paradigm participants would only be shown a few objects, often merely two. After target offset, but before the memory test, participants would be informed which of the two objects they would have to recall. This allowed them to prioritize one object in memory over the other and stabilizing that object, concentrating the limited resources of working memory, which may have initially been more distributed across multiple objects. When then asked to report this object, participants would already have it in the focus of attention, and its representation strengthened. The primary application of this paradigm usually is the study of “internal shifts of attention”, the retrospective prioritization of objects in memory. One dominant hypothesis to explain the retro-cue effect is that this prioritization protects the object representation from interference by the memory probe itself. In change-detection experiments (a typical memory paradigm), participants are shown a stimulus and asked to report whether it was the same or different from the one they previously saw and memorized. If that memory representation was very weak, the presentation of the probe item may simply override the memory of the old object. The participant would be left with no way to compare the previous and current percepts as the previous percept was lost in the void of time and neural noise. In fact, the process of retrieving a memory representation itself might be destructive to the representation, even in other paradigms than change-detection. Is it possible that common paradigms underestimate the true capacity of working memory because they do not account for memory representations too weak to survive the recall process?

One of the two labs that concurrently established the retro-cue paradigm already proposed that the retro-cue effect may indicate that there is more to working memory than traditional estimation methods not using retro-cues suggest (Landman et al., 2003). Sligte et al. (2008) went a step further and showed that indeed, with the use of retro-cues, working memory capacity estimates can rise from 4 (out of 32 initially presented objects) to 16! Sligte et al. coined the term

“fragile memory”, to denote a form of short-term memory conceptually situated between iconic memory and working memory in terms of longevity, stability and capacity. Fragile memory is detectable for several seconds (as opposed to only a few hundred milliseconds as is the case with iconic memory) and has a considerably larger capacity than working memory. As the name suggests, it is very susceptible to visual interference. However, fragile memory can only be disrupted by visual information resembling that of the initial memory trace, while iconic memory can be destroyed even by a light flash. This implies that fragile memory representations consist of retrievable objects, rather than mere low-level artefacts of the visual processing stream. Importantly, fragile memory representations could be formed without receiving a great deal of attentional resources (which may pose a fundamental limit of working memory capacity). A more complete discussion of the similarities and differences between these memory forms can be found in the re-prints in sections 2.2 and 2.3.

Considering the degree of entrenchment of the “traditional” models of short-term memory (e.g. Baddeley & Hitch, 1974), the proposal of a third form of visual short-term memory was bold, but not unreasonable, given the experimental data. Surprisingly, this finding has never been incorporated into current models of short-term memory. Several alternative explanations and critiques were proposed, but neither was able to really account for the observed dramatically high-capacity estimates. The most viable alternative explanation is that participants grouped the visual display into clusters of objects (chunks) and in this way used a more efficient memorization process. However, even if that is the case, it changes nothing about the amount of *information* that is stored and retrieved in this paradigm, superseding traditional estimates of working memory capacity. This finding may, for example, indicate efficient compression mechanisms of working memory storage that allow for more information to be stored in more naturalistic, real-life settings than during a typical lab experiment. To this day, Sligte et al.’s proposal has received very little attention and no attempts of reconciling his findings with the literature have been made. Very slowly the field of working memory research is sometimes moving beyond impoverished stimulus displays which are treating “working memory objects” in isolation. For example, investigations into ensemble perception and interactions between concurrently held memory representations are beginning to

address the gap between the richness of perception and the “perspective of the behavioural scientist”. If nothing else, fragile memory and the retro-cue effect clearly demonstrate that memory capacity is very much a function of the employed paradigm, rather than an objective, task-independent measure of a cognitive function. It is very possible that short-term memory exists on a continuum between very fragile representations with a near unlimited capacity (perhaps including perception itself) and ultimately a single, very reliably represented object with which an observer can perform an action.

In the second year of my PhD, I did not know much about these intricacies, but sensed that there would be great value in taking a closer look at the retro-cue paradigm and fragile memory. From my work on spatial remapping (see section 1.1), I knew that it was commonly assumed that only very few objects could be remapped, namely those objects that were attended (prioritized) and placed in robust working memory. This opened up an interesting question: could fragile memory representations survive the necessary computations required to update their locations across eye movements? If so, this would imply that remapping was not limited to the most robustly encoded items. It would also provide evidence for the idea that fragile memory representations were accessible and actively maintained. My approach was simple: participants memorized a rather large number of orientated lines and were asked to recall one of these items after having made an eye movement. In line with previous findings, my results indicated that on average, only a few of these items (three or four), could be recalled. However, when presented with a retro-cue prior to the memory test, but after the eye movement, participants were able to recall more items as compared to a condition without a retro-cue. This was true for conditions with and without eye movements and clearly demonstrated that even memory representations which were initially unattended (or less attended; and thus fragile) were nonetheless updated during eye movements. If fragile memory was not updated across eye movements, memory performance should have been the same with or without a retro-cue.

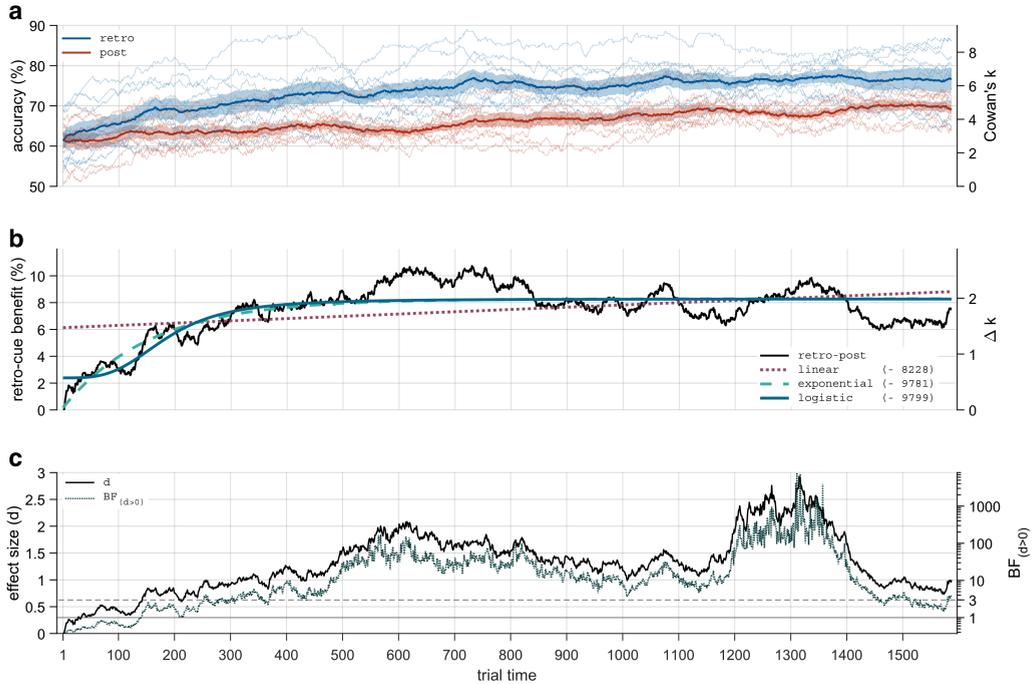
One very nice aspect of this project was that it gave me a chance to apply my new-found appreciation for Bayesian statistics and linear mixed effects models (LME). Bayesianism is an increasingly popular

alternative to the traditional frequentist statistics with numerous benefits such as the emphasis on generative modelling, the ability to quantify evidence using Bayes factors, allowing the use of optional stopping procedures (collecting data until an evidence threshold is reached rather than having to rely on a pre-determined sample size), and allowing to differentiate between evidence for a hypothesis, against a hypothesis and inconclusive evidence for a hypothesis. A second important addition to my statistical method repertoire was the use of LMEs. These are essentially regression models that offer at least two important advantages. In contrast with other generalized linear modelling (GLM) approaches (e.g. ANOVA), they provide a unified, intuitive and structured approach to pretty much any and all statistical analyses a cognitive scientist would ever need. Most importantly, LMEs differentiate between different types of sources of variance. For example, while a GLM makes the assumption that a manipulation in a behavioural experiment has the same effect on all participants, LMEs can model an effect for each participant separately. This has the obvious advantage of separating variance induced by individual differences from variance induced by the experimental manipulation, making these techniques much more powerful. LMEs could also be performed in a Bayesian framework, and in fact, are very well suited for this. Thankfully, I did not have to do all the heavy lifting of the implementation, as the legendary Paul Bürkner had produced an R package called brms, which allowed the implementation of Bayesian LME models using the standardized Wilkinson notation (e.g., $\text{mpg} \sim \text{weight} + \text{acceleration}$, meaning: miles per gallon predicted by the variables weight and acceleration). This made it very straight-forward to implement statistical tests. Of course, there is much more to statistics than just running a script, and so, in order to make sure I get things right, I invited Kees Mulder from the statistics department at UU to collaborate with us on this paper. During many work sessions we sat together, and he patiently explained all the intricacies of the analysis. This had two major advantages: I would be sure that my analyses were correct, and I would learn in much greater depth how to perform these types of statistics in the future. Kees was very nice to work with and as a statistician was also happy to experience the behavioural scientist side of things. I can heartily recommend collaborating with statisticians, as oftentimes the nuances of analyses can be very tricky to get right, and it would simply be impossible for a

INTRODUCTION

researcher to operate on the high level of a dedicated statistician. Kees became a co-author and the project was eventually published in Scientific Reports (section 2.2).

1.4 120 hours of oriented rectangles



It is possible, that fragile memory is not a mechanistically different form of short-term memory. Short-term memory items could exist on a continuum of stability within the same system, dependent on the amount of limited memory resources (e.g., attention) dedicated to each item. In that case, the retro-cue manipulation would simply increase the resources available to an initially weakly attended (and weakly represented) item and stabilize it enough to become detectable in a working memory paradigm. Alternatively, retro-cues might (as for example suggested by Sligte et al.) pull information from a less resource-limited (but implicit) memory system and place it into reportable working memory. This latter possibility made a very clear prediction: if participants could indeed store a large amount of information in a form of short-term memory that became accessible through the deployment of a retro-cue, then capacity estimates should only be limited by the ability to make use of that retro-cue. Given

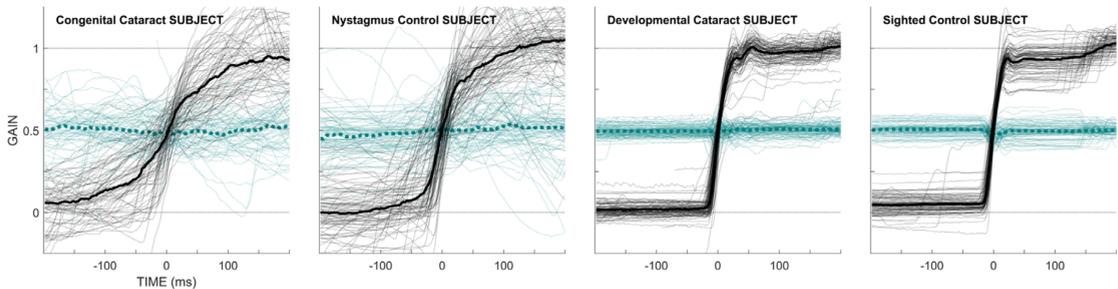
enough practice, capacity estimates should rise to a very high number, well beyond traditional estimates of working memory capacity (the condition without retro-cues, which was expected to plateau very quickly). To test this prediction, 10 participants came to the lab for a total of eight consecutive sessions, during each of which they performed a change-detection task with a set size of 12 oriented rectangles, which they had to memorize and later compare to a test rectangle. In half of the trials a retro-cue would indicate the location of the test item before it would appear. This meant that each participant performed the task for 12 hours (quite frankly, a heroic deed, easily appreciated by anyone who spent any amount of time performing a psychophysical task). A comparison of memory performance over time between the conditions with and without a retro-cue would reveal whether or not a qualitatively different memory system was accessed through the use of retro-cues. The results revealed that the retro-cue conveyed a consistent benefit to memory performance, but the difference between conditions plateaued quickly and did not continue to rise. This was very clear evidence that retro-cues did not access a resource-unlimited, separate form of memory, but simply shifted memory resource between items stored within the same memory system.

My findings have two important implications for working memory capacity estimation: First, there is more to working memory than can be detected in pure reproduction or change-detection tasks. For decades, investigations using traditional working memory paradigms (without retro-cues) may have underestimated participants' working memory capacity! This is particularly striking as working memory capacity is associated with a wide range of cognitive aspects, including reading ability and intelligence. Of course, this does not mean that all of this research is flawed, especially if capacity estimates were not considered to be an absolute, relative measure. When, for example, comparing the measured capacity to other participants in the same study using the same paradigm or other experimental conditions the relative nature of working memory capacity does not pose much of a problem. However, this is not always the case. My findings demonstrated once again that it is not the "true" capacity of working memory that is measured in lab experiments, but the capacity detectable by the employed paradigm. The true capacity of human working memory to this day is unknown. We can only ever

measure the task-specific performance of an individual under specific circumstances.

Secondly, an additional, unexpected finding was that memory performance continued to increase in both conditions! While the measurement of interest in my study was the difference between conditions, the unexpected linear increase in performance in both conditions which continued until the last session revealed an uncomfortable fact: working memory capacity, a supposedly stable trait underlying many cognitive functions was not stable at all. I have all reason to assume that the measured capacity would have continued to increase if the experiment sessions would have continued. If working memory capacity is not a stable trait, but dependent on the familiarity of the participant with the task, what was really being measured? How could memory capacity be measured and reliably compared between individuals? Especially when different individuals learn at different speeds, this could easily confound results in correlational studies. Similarly, different studies typically employ different amounts of practice trials, making meta-reviews comparing different studies less reliable. More importantly, this finding revealed once again that the supposedly objective estimation of memory capacity determined in a lab experiment measures a construct, which at best correlates with a real biological mechanism and vaguely reflects an aspect of cognitive function, but which should never be confused with the real thing. Keeping in mind these intricacies of epistemological reasoning, our results were ultimately published in *Memory & Cognition* (see section 2.3).

1.5 Eye movements in the formally blind



On a particularly fateful day at Utrecht University, a visiting German researcher, Prof. Brigitte Röder, gave a lecture at our institute, and as it was the custom at the time, a few PhD candidates went to lunch with her. She told us about her project in India. As I had travelled there before, I expressed my love for this country and we talked about India, about vision science and what it means to complete a successful PhD. To my great shock, a few days later I received an email with an offer to come work for her. After some bureaucratic juggling we arranged that I could extend my PhD time and join the project for a year, which later became two years, increasing my total PhD time to six years. I left my life in Utrecht behind and ventured to Hamburg to work on an incredibly unique and challenging project.

The mission was to contribute my expertise in eye tracking to an ongoing collaboration with the LV Prasad Eye Institute (LVPEI) in Hyderabad, India. This hospital and research institute is performing extremely impressive work in the realm of patient treatments of all kinds of vision-related impairments, from myopia to corneal transplants, as well as various lines of high-profile research reaching from ophthalmology to genetics. What probably impressed me the most was that the LVPEI offers free treatment to patients who cannot afford it, and that without having to prove their socioeconomic status. Someone might, for example, come to the LVPEI and say that her father has developed cataracts in both eyes but just cannot afford the operation. Her father's vision will be restored by means of a cataract operation and the implantation of an artificial lens, at no cost. There

were other signs of progressive thinking at LVPEI, such as men and women, researchers and cleaners, nurses and doctors eating together in the same canteen (something absolutely uncommon in India's caste system).

The scientific mission of the Hamburg-Hyderabad collaboration involved individuals who were born with complete bilateral cataracts. Such individuals are legally and practically blind. While their retinas are typically intact, the opaque lenses prevent any patterned visual input to reach their visual system. In western Europe babies born with the characteristic milky lenses would be operated very shortly after birth. However, India's vast rural landscape does not always allow for such treatment, both for financial reasons as well as simply a lack of information. The LVPEI was founded by a prominent figure of Hyderabad's film Industry, L. V. Prasad himself. Being a movie director and patron of film, he realized that many people are in fact unable to see movies or anything else for that matter. He decided to fund an eye hospital that would bring vision to as many people as possible. Ambassadors from LVPEI regularly venture out to remote locations in Southern India to identify cataract cases, for example in rural schools for the blind, and inform these people that they do not have to be, in fact, blind. They are invited to come to Hyderabad or one of LVPEI's sister locations in order to get a cataract operation. It should be noted that well before LVPEI started doing this, the first project of this kind was initiated by Prof. Pawan Sinha in New Delhi, in Northern India, under the name Project Prakash.

Individuals who were born with complete bilateral cataracts not treated immediately after birth share one unique trait: their brains did not receive visual input during a sensitive period of the development of the visual system. In many cases, it was many years before the operation. This is very meaningful to researchers interested in brain development. Do visual system functions still develop if the first visual input was received later in life? In other words, do certain visual system functions critically rely on that sensitive period during the first months of life? The Hamburg lab had already published research on this patient group in the realm of face recognition and multi-sensory perception using EEG and behavioural tasks. I came in as an eye movement specialist and my task was to develop and execute experiments that sought to investigate the formation of the eye

movement system in individuals who regained vision following the cataract operation. For the first few months I stayed at Hamburg University and developed a series of experiments that would examine basic eye movements in this patient group, as well as some aspects of visual working memory. A major challenge was the known presence of nystagmus in cataract-reversal individuals. Clinical nystagmus is the inability to fixate an object, that is, to keep the eyes still. Nystagmus patients' eye movements look very different from healthy individuals. Their eyes dart back and forth multiple times per second in an oscillating saw-tooth pattern, continuously circling an object rather than fixating it. However, eye trackers rely on stable fixations in order to calibrate the system, i.e., to map eye camera position signals to screen coordinates. Thus, my first task was to develop a manual calibration system, with which a researcher would examine raw eye tracker data traces to identify (that is, to make a wild guess) where on the screen the nystagmus patient was looking when a calibration target appeared. While this served as a rough calibration estimate, I later refined this process by leveraging the endpoints of eye movements (which seemed to briefly fixate a visual target, before drifting off again) in an iterative optimization process: define the eye positions just after target onset and use these endpoints to calibrate the eye tracker, then use that calibration to re-define eye movement endpoints, use these endpoints to re-calibrate, etc. Repeating this procedure many times with many eye movements all across the screen eventually zeroed in reasonable calibration coefficients. While this procedure would not allow for the eye tracker precision I was used to in my experiments with healthy individuals, it eventually allowed, for the first time, to investigate goal-directed eye movements in cataract-reversal individuals with severe nystagmus.

Before I could run any of the experiments on this patient group, I of course had to fly to Hyderabad, where I ended up staying for about eight months in total. Sometime prior, Prof. Karl Gegenfurtner had gifted an EyeLink1000 eye tracker to the LVPEI, which, however, had not seen a lot of use since then. One of my first tasks was to establish an eye tracking lab. While I was familiar with the method of setting up the EyeLink1000 system from my time at Utrecht University, the reality of setting it up in an Indian eye hospital was a whole different beast altogether. What one has to consider is that life and work functions differently in India. Having visited India twice before for the

purpose of unhinged and extended travel, working there was very different. More than ever before I had to think on the fly and improvise wildly. I was absolutely impressed by the efficiency and effectiveness of the LVPEI, by Indian or any standards. The staff was generally very helpful and accommodating, especially the glowing examples of Dr. Ramesh Kekunnaya, Emeritus Prof. Dr. Balasubramanian (popularly known as Dr. Balu) and Idris Shareef. An experiment room as assigned to us in the hallway of the paediatrics department, which made sense, since most of our participants were children. However, this also meant that there was an incredible amount of activity going on at any given time, often involving crying children in the next room. Due to the large volume of patients being seen sometimes the eye tracking room itself was also used as examination room in between experiments. I would often be informed at very short-notice that a potential participant had arrived and rushed over to perform the experiment. Since the children and family rarely spoke English, I had to find a nurse or doctor to volunteer to translate, which meant I had to first explain the experiment to my translator and hope that these instructions were conveyed accurately. Of course, I was not entirely left alone in this. Idris Shareef and later also Siddhart Rajendran assisted with the data collection, and in turn I sometimes helped out with the other behavioural and EEG experiments.

Setting up the lab and eye tracking computers in Utrecht was mostly handled by dedicated IT services. In Hyderabad, when I required an extra keyboard or HDMI cable this meant that I had to venture into Hyderabad's computer district, which involved surviving crossing the road, finding a rickshaw, haggling over the price, surviving the half hour drive through insanely chaotic traffic, and then navigating the maze of hundreds of little shops offering everything from Windows-95 era computers to modern iPhones. It is very possible to find what one is looking for eventually, but it might and very likely will take all day. Similarly, appointments with our participants were not made for a specific time, but e.g., for next Wednesday. It should be noted that LVPEI doctors were actually very reliable when it was within their power (on some days there were so many patients that waiting times would average several hours). Having lived and travelled extensively in India before, it was clear to me that given the circumstances of socioeconomic and every-day reality of a rapidly growing country with a population size of over 1.6

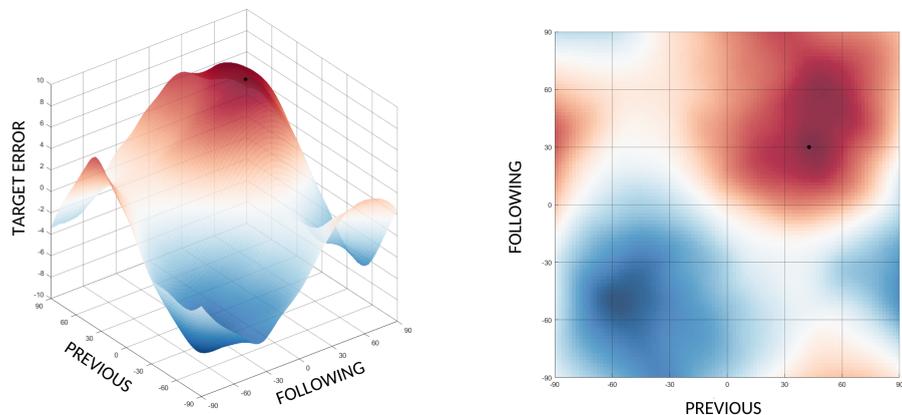
billion human beings, the climate, the disparity between deeply rooted traditional systems and the drives and demands of the modern world, it is absolutely impressive how well Indian society functions. Uncertainty is accepted as part of life. Things happen. Sure, you try your best to get to a place at a certain time, but this is often humanly impossible. I have observed many times during my travels that “Westerners” thrown into these circumstances are unable to cope, resulting in anger and frustration. An Indian driver on the other hand is not prone to road rage, because if she were, she would be yelling at all times. Instead, traffic in India is characterized by a certain stoicism amid the hectic exterior, including the constant honking (which is an expression of everything from “hello” to “I’m going to overtake you now”), a total disregard for any lane markings (even if present) or right-of-way rules. One observes the person in front, behind, left, and right and acts accordingly to the free-for-all flow of traffic. Of course, this is only possible in combination with the acquisition of super-human ninja skills of reaction time and spatial awareness. The same goes for crossing a busy road. It’s quite literally playing a sort of faith-based frogger. Faith in the idea that if you do not hesitate but bashfully step one lane further, the upcoming car will see you and stop, even though while approaching it will give absolutely no sign of this until you actually move. The general advice is to find someone, for example an Indian grandma, who is crossing the road and follow her example.

These are just some examples of the unique experience that is life and work in India. Of course, I was privileged as an outsider from a comparatively wealthy country, to not endure the full depth of all of India’s realities and I have the deepest respect for all Indians for being able to thrive in the circumstances of their society. I want to be absolutely clear: despite the culture shock experienced by a privileged white boy like me and having spent almost two years of my life there, I am in love with this country. Despite many shortcomings (especially within the realm of human rights for some sections of society), this ancient culture has much to teach, and to delight us with. The food, for one, is absolutely phenomenal. There are no words to describe how much I enjoyed it, every single time. Curries cooked for hours, according to recipes handed down from mother to daughter over generations, made from local vegetables grown in the marvellous Indian soil. As cooling trucks are not at all common, most produce comes from nearby farming communities. I also experienced a special

kind of gentleness in India, one that perhaps isn't too apparent on the harsh surface but does emerge when it comes to dealing with conflict and with suffering. I also experienced Indian people as deeply sentimental and strongly embedded in their many cultures and religions, which permeate all facets of daily life, lived fully, with a power, a life energy that I miss in the colder parts of the world.

I eventually concluded the research, demonstrating for the first time that cataract-reversal individuals are able to make visually-guided eye movements that were not worse than those made by other individuals with clinical nystagmus. The manuscript was eventually published in the *Journal of Vision* and is re-printed in section 2.4.

1.6 Serial dependence



Before I left for Hamburg and India, I had piloted a series of experiments, in which I aimed to determine how stable different forms of visual working memory are. I approached this through “masking”. Masking refers to any disturbance that is introduced in addition to a stimulus. In my case, I would present visual stimuli, which participants had to memorize. I would then display interfering visual displays, such as noise or stimuli that were highly similar to the memory targets in an effort to disrupt the encoded representation of the memorized targets. To my surprise, this turned out to be extremely difficult. Whatever I threw at participants in terms of masks, none would produce more than a few measly percent of performance reduction. Even when I displayed a barrage of stimuli that included a target that had to be memorized in addition to the initial memory targets, this additional attentional diversion did not impact participants’ memories by much. Visual working memory turned out to be incredibly robust to visual interference, even when participants had to memorize many targets simultaneously.

Perhaps my use of a change detection paradigm, in which the to-be-memorized line stimuli could have one of two or one of four orientations was not sensitive to pick up on the distortions created by masks. I then started to wonder whether it would not be better to try to coax the visual system into updating or replacing memory

representations, rather than attempting to destroy them. It was around this time that I went on a hiking trip to Iceland, towards the end of which I paid a visit to the University of Iceland and the lab of Arni Kristiansson, the Icelandic Vision Laboratory. I considered that since I would probably not often be in Reykjavik, I could use this opportunity to say hello to the research group there, which I had regularly encountered online but never in person. As I did so, Arni informed me that there were some other researchers currently visiting and decided to make an impromptu vision-science symposium. COVID-19 case numbers were very low in Iceland at the time and so it was possible to hold what we would later call the only in-person vision science conference of 2020. I gave a talk about my previous work, and during socializing chatted about my current attempts at working memory distortion and updating. It was suggested to me that what I was trying to do sounded a lot like “serial dependence” but in reverse. I had briefly heard about serial dependence before, in the context of a nuisance effect in experiments with sequences of trials. As it turned out, almost all of the other talks at the symposium were in fact about serial dependence and about its influence on behavioural reports, not as an artefact, but as a meaningful cognitive process in its own right. Serial dependence is a type of visual history effect, in which previously observed information influences (biases) the perception or encoding of currently observed stimuli. Conversely, what I wanted to do was influence previously memorized stimuli through currently perceived stimuli: serial dependence reversed in time. A light dawned on me.

Once I returned from Iceland I got to work and started to implement an experiment consisting of a sequence of orientations, which participants had to memorize. They were then cued to reproduce one of the four orientations. If my idea was right, then not only previously observed orientations should influence the reported orientation of targets later in the sequence, but orientations presented after the target was perceived, encoded and no longer visible should also influence behavioural reports (for example, when the second orientation had to be reproduced it should be influenced not only by the first, but also by the third and fourth orientations). And this is exactly what I found. Starting from this initial finding, I conducted a number of follow-up experiments, which can be found in the re-print of section 2.5.

Beyond this scientific breakthrough, this project taught me another important lesson: the perfectionism with which I had approached lab-based experiments was perhaps not at all necessary. As these were the times of the great plague of 2020, I was no longer able to run experiments in the lab. Instead, I learned javascript and created an online version of the experiment using jsPsych, a toolbox for creating psychophysical experiments in javascript. The experiment was then hosted on a webserver using the JATOS experiment server software. Participants were recruited via the Prolific web service. Not only did this method deliver results that were just as good as what could be obtained in the lab, it also freed up experimenter time. I was able to collect data from potentially hundreds of participants in merely an hour or two of releasing the experiment on the web. While collecting data from a few dozen participants in the lab took many weeks and each hour of experiment time would require an hour of experimenter time, handholding participants in the darkness of the lab, I could now test hundreds of participants in parallel. These participants were doing the experiments from the comfort of their homes, which meant that screen sizes and lighting conditions differed, presentation times were not exact on a millisecond level, and I had no control over potential distractions in the visual field of the participants. Despite this, the rate of “bad responders”, aka random-button pushers, was not much higher than that obtained in lab experiments and the data quality very adequate. Were we overthinking lab-based experiments this whole time? Many people I had talked to had made similar experiences and came to similar conclusions. In the future, there would surely be use cases for highly-controlled lab experiments (for example when colour of luminance needs to be tightly controlled), but it had become clear that in the future many behavioural experiments would be taking place online, even after the pandemic.

As I was analysing the data of these experiments, I came across a job ad for a post-doc position, offered by no one else but Prof. Arni Kristjansson, the head of the Icelandic Vision Laboratory, together with another researcher, David Pascucci from Genova, Switzerland. The topic of the project was serial dependence and related concepts. Of course, I applied immediately and during the interview was able to show my current work as it was extremely relevant to the offered position. About a month later I received word that I was accepted. I

INTRODUCTION

would be a post-doc in Reykjavik, Iceland! The lab that gave me the idea for my current research hired me to continue this exact line of work for another two years. A full circle.

1.7 Epilogue

I am writing these words from the middle of the Atlantic Ocean on a three-day ship journey that will bring me from the northern tip of Denmark to the eastern coast of Iceland to start my new life. It had been a long journey spanning several continents. Born in Kyrgyzstan, my family moved to Germany when I was two years old, where I grew up and went to school. I escaped to Groningen, Netherlands when I was 19 years old in order to study Psychology, which included an exchange semester in Guildford, UK. I then returned to Groningen to finish my bachelor's degree. After a subsequent, long trip to India, I began my neuroscience research master studies in Utrecht. I did not stay for long and went to Nijmegen for a half year internship. After returning to Utrecht and conducting my PhD research for two years I moved to Hamburg, from which I continued to Hyderabad, India to implement experiments and collect data, before returning to Hamburg to write up and publish the resulting research.

Returning to Utrecht in order to finish the remaining two years of my PhD proved somewhat challenging as mother nature had decided to unleash the plague on all our houses, and after just two months of joyous reunion with my colleagues and the city of Utrecht, which I had come to love, life was cancelled, and we all started working from home. Just before the borders closed, I decided to flee to Brussels, Belgium to be with my girlfriend of the time rather than to remain isolated. In Brussels I continued my work as I had in so many places before. At least, I did as well as I could under the circumstances that we all faced at the time. After a few months we found an apartment in Utrecht and moved back there. Finally giving in to the strain of the pandemic, our relationship, not entirely unexpectedly, came to a conclusion on the edge of the Sahara near Mhamid, Morocco during a trip we took to escape the madness our lives had become, in an attempt to save our minds. I decided to stay in the desert for a few weeks longer to deal with my broken heart. I set up a desert office in a rather comfortable clay hut with the great sand dunes in my backyard. This went rather well overall, although surprisingly, high-speed internet was far easier to come by than a chair. As I had nothing much left in Utrecht with the pandemic still raging and city and university life shut down, I decided

to move to the Belgian countryside near Ronse, with a group of friends I knew from days past and to continue work from there.

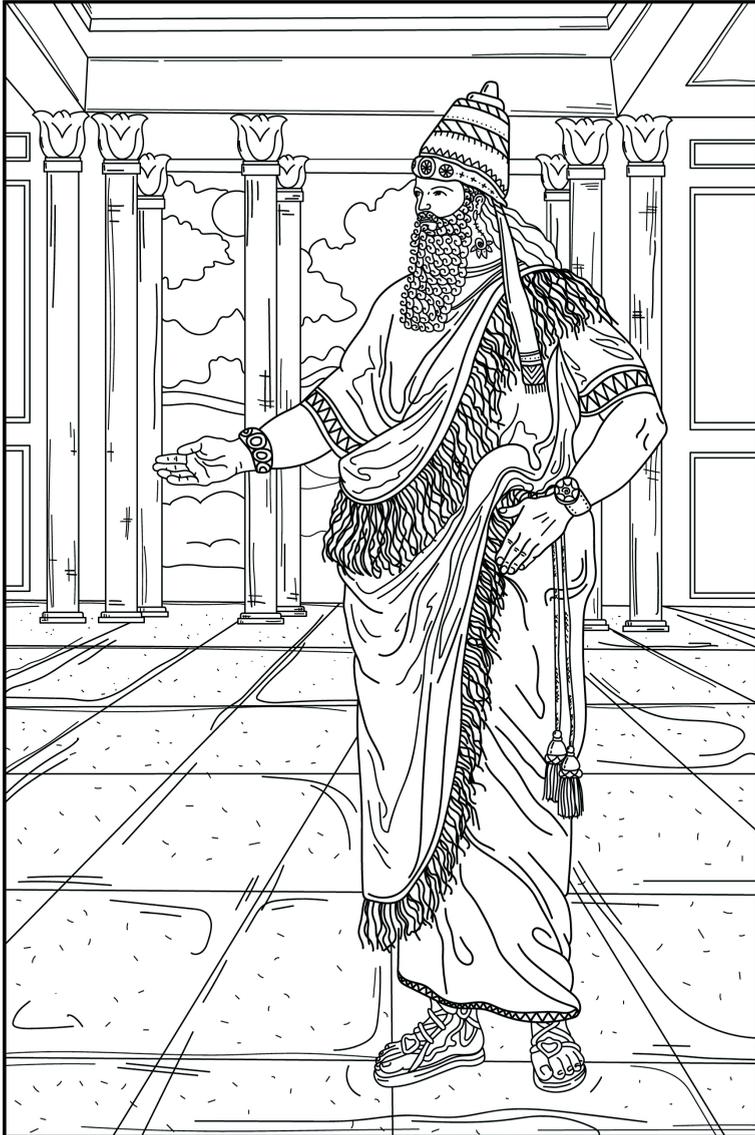
Finally, and towards the end of the PhD contract I gave up on the concept of home entirely, bought a car and drove towards the Swiss alps. My mission was to write the last chapters of this thesis while immersed in the nature that I loved so dearly, the forests and mountains of central Europe. If I had to work remotely and in relative isolation, I could do this from anywhere in the world. I had built a second battery, power converter and sound system into my car, in which I could also sleep very comfortably. I parked the car in various places, in peaceful remote valleys and on mountain tops with a view towards even greater, gorgeous giants, befriending shepherds, herding sheep, hiking, writing, and thinking.

As I am now scribbling these lines, we lay on anchor at Tórshavn, the major port of the Faroe Islands, about a day by sea from Iceland itself. Tomorrow we moor in Seyðisfjörður, from which I will drive all the way around the island of Iceland to Reykjavik, where I will begin my new life as a post-doctoral researcher. Three days from now. As I close this chapter of my life by submitting these pages, I am ready to begin the next.

1.8 References

- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *The Bell System Technical Journal*, 40(1), 309-328.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press
- Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience?. *Trends in cognitive sciences*, 20(5), 324-335.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of cognitive neuroscience*, 15(8), 1176-1194.
- Lamme, V. A. (2010). How neuroscience will change our view on consciousness. *Cognitive neuroscience*, 1(3), 204-220.
- Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision research*, 43(2), 149-164.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied*, 74(11), 1.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores?. *PLOS one*, 3(2), e1699.

CHAPTER 2: RESEARCH ARTICLES



2.1 Error compensation in random vector double step saccades with and without global adaptation

ABSTRACT

In saccade sequences without visual feedback endpoint errors pose a problem for subsequent saccades. Accurate error compensation has previously been demonstrated in double step saccades (DSS) and is thought to rely on a copy of the saccade motor vector. However, these studies typically use fixed target vectors on each trial, calling into question the generalizability of the findings due to the high stimulus predictability. We present a random walk DSS paradigm (random target vector amplitudes and directions) to provide a more complete, realistic and generalizable description of error compensation in saccade sequences. We regressed the vector between the endpoint of the second saccade and the endpoint of a hypothetical second saccade that does not take first saccade error into account on the ideal compensation vector. This provides a direct and complete estimation of error compensation in DSS. We observed error compensation with varying stimulus displays that was comparable to previous findings. We also employed this paradigm to extend experiments that showed accurate compensation for systematic undershoots after specific-vector saccade adaptation. Utilizing the random walk paradigm for saccade adaptation by Rolfs et al. (2010) together with our random walk DSS paradigm we now also demonstrate transfer of adaptation from reactive to memory guided saccades for global saccade adaptation. We developed a new, generalizable DSS paradigm with unpredictable stimuli and successfully employed it to verify, replicate and extend previous findings, demonstrating that endpoint errors are compensated for saccades in all directions and variable amplitudes.

Published as:

Zerr, P., Thakkar, K. N., Uzunbajakau, S., & Van der Stigchel, S. (2016). Error compensation in random vector double step saccades with and without global adaptation. *Vision research*, 127, 141-151.

All authors designed the study concept. PZ programmed the experiment, collected experiment data, and wrote the manuscript. PZ and SU conducted the analyses. Critical revisions were provided by all co-authors.

INTRODUCTION

When observing the visual world we typically scan the environment with successive fast eye movements (saccades). This brings relevant objects onto the fovea, an area at the center of the retina specialized for color and sharp vision (e.g. Provis et al., 2013). For each of these saccades a motor vector is calculated based on current eye position and the saccade target, which requires gaze centered (retinotopic) spatial maps to be maintained and continuously revised. To update the representation of a retinotopic location after a saccade the visual system can rely either on visual input or, in its absence, on a positional recalculation based on spatial working memory and knowledge of eye displacement (remapping). Predictive remapping even before an eye movement is executed (Duhamel et al., 1992) facilitates rapid programming of successive saccades. Predictive updating of receptive fields has been observed in neuronal populations in the lateral intraparietal area (LIP) (e.g. Duhamel et al., 1992, Medendorp et al., 2003), superior colliculus (SC), extrastriate cortex, and frontal eye fields (FEF) (reviewed in Colby & Goldberg, 1999).

The double step saccade (DSS) task is commonly used to study spatial remapping in the absence of visual feedback (Hallet & Livingstone, 1976). In this paradigm two targets (**T1** and **T2**) are flashed briefly, and participants saccade from the fixation target (**F**) to T1 (first saccade; **S1**) and from T1 to T2 (second saccade; **S2**). To successfully land on T2, its fovea-relative location after the saccade needs to be recalculated based on memory of its last known retinotopic position and the saccade vector to T1.

Errors in saccade landing positions create a mismatch between the ideal and actually executed motor vector. A motor vector is defined here as the actual saccadic motor output as measured by the eye tracker. An ideal saccade is defined as the vector between the endpoint of the last saccade and the position of the current saccade target. In order to correctly remap a future target, landing errors of the current saccade need to be accounted for. It has been suggested that a copy of the saccade motor vector (corollary discharge; **CD**) supplies eye displacement information necessary for spatial remapping (e.g. Guthrie, Porter & Sparks, 1983; Sommer & Wurtz, 2006). Collins, Rolfs, Deubel & Cavanagh (2009) observed that CD would closely represent the actual saccade and thus incorporate trial-by-trial saccade error. Joiner, FitzGibbon & Wurtz (2010) recently reported near perfect compensation in second saccades for endpoint errors in

first saccades. They observed a strong correlation between first saccade error and compensatory component of the second saccade in the opposite direction, a finding that is best explained by the availability of a CD vector containing saccade errors to the visual system.

Further support for the reliance of remapping on CD comes from studies of saccadic adaptation. In the intra-saccadic back-step task originally developed by McLaughlin (1967), participants make saccades to targets that are displaced slightly towards the previous fixation target as soon as a saccade is detected. Due to saccadic suppression, subjects are normally unaware of target displacement and instead the visual system detects apparent systematic overshoots and gradually shortens saccades (for reviews of saccade adaptation see Hopp & Fuchs 2004; Pélisson et al., 2010; Iwamoto & Kaku, 2010; Herman et al., 2013). Tanaka (2003) examined remapping in monkeys with a DSS task using horizontal-vertical saccade pairs while gradually inducing adaptation in first saccades. By regressing the compensatory (horizontal) component of second saccades on the horizontal endpoint error of first saccades, he found that the monkeys were able to compensate for gain changes in first saccades to correctly land on T2 (82% and 85% respectively for the two monkeys used in the study). He concluded that at least half of the signal corresponding to a CD vector must have come from sites downstream from the site of adaptation (and therefore included information about the change in saccade amplitude). However, because induction of adaptation and probing of compensation occurred gradually and simultaneously, natural trial-by-trial error and variations in gain induced by adaptation were treated equally while these errors could potentially be represented differently in the visual system. Collins (2010) examined error compensation in a similar way in humans by again regressing the (compensatory) horizontal component of second saccades on the horizontal amplitude errors of first saccades. In contrast to Tanaka, Collins observed saccades before and after inducing adaptation. Identical regression slopes were observed in adapted and unadapted saccades (corresponding to around 74% compensation on average in both cases), indicating that the visual system is indeed aware of and compensates for systematic changes in saccade amplitudes. Despite the difference in amount of compensation reported these studies agree that first saccade endpoint errors are mostly accounted for in subsequent saccades.

The vast majority of the remapping literature depends on stimuli with fixed and limited target positions (fig.2A), which simplify analysis but may be less indicative of natural behavior and could create experimental confounds if subjects memorize the small set of stimulus locations. If this were the case, the accurate error compensation observed in these studies could be explained by a spatiotopic long-term memory representation of target locations rather than a reliance on trial-by-trial information from CD. Zimmermann (2013) reported spatiotopic displacement of a saccade target following saccade adaptation, indicating the potential of long-term spatiotopic memory of target locations after repeated exposure to the same stimuli. The global saccadic adaptation paradigm by Rolfs, Knapen & Cavanagh (2010) liberated experiments that examine the effect of adaptation on spatial remapping from this restriction. In their adaptation task targets appeared at random positions on the screen in each trial, and saccades were successfully adapted for all directions (parametric adaptation). This provided a means for studying saccade adaptation as a global change in the saccade system. Our study aimed to extend previous studies that examined saccades in the DSS paradigm with and without adaptation by utilizing highly variable (and thus unpredictable) stimulus displays. In the first experiment we first introduce a novel DSS paradigm with random target vectors (fig. 2B) and accompanying analysis method (fig. 4) as well as models using simulating data to validate this method (fig. 5). We then examined error compensation in a slow-paced (high pre-programming) and fast-paced (low pre-programming) DSS task. Ditterich et al. (1998) suggested that high pre-programming should reduce error compensation in sequences because saccades are then executed based on fixed amplitudes. In this case better error compensation should be observed in our fast condition. In a second experiment we compared the amount of error compensation in second saccades before and after *global* saccade adaptation in the context of random target vectors. If, as has been suggested, CD contains adapted motor vectors then we should observe appropriate error compensation after adaptation.

It is so far unknown to what extent the mechanism for global adaptation differs from that of specific vector adaptation. Until recently the mechanism behind saccadic adaptation has been thought to be highly selective for the adapted saccade vector in terms of position, direction and amplitude as well as for saccade type (e.g. reactive versus memory guided). Transfer to other vectors was found to be very limited (Deubel, 1995; Noto et al., 1999) and results for transfer to other saccade types has been mixed (for an overview see

Péllisson, 2010, Kojima et al., 2015). The existence of context specific adaptation beyond specificity to particular vectors demonstrates that saccade adaptation is not just a simple motor recalibration process depending on motor states such as orbital eccentricity but can use predictive visual cues to change saccade metrics depending on visual properties of saccade targets (Azadi & Harwood, 2014). For instance, Herman et al. (2009) found specific adaptation to flickering versus non-flickering stimuli. However, these results have usually been obtained by adapting only one or a few specific vectors. Rolfs et al. (2010) demonstrated that global adaptation for all saccade directions and amplitudes is possible when displaying targets in a random walk design instead of using fixed vector positions. Importantly, they demonstrated that saccades were not simply adapted for many different vectors but that this adaptation was *parametric*: a global change in the saccade system. We examined for the first time whether this type of adaptation also transfers between saccade types. If transfer of global adaptation is limited to saccade type, we would expect that adaptation for the second (memory guided) saccades in the DSS task is reduced or absent as they are qualitatively different from the adapted (reactive) saccades. Conversely, if global adaptation is parametric beyond amplitude and direction both saccades should be adapted.

MATERIALS AND METHODS

This study is approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University and has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Participants

Experiment 1 10 Utrecht University students (mean age: 22.1 ± 3.7) participated in the experiment and were rewarded with EUR 6,- per hour. All participants had normal or corrected to normal visual acuity and gave informed consent.

Experiment 2 20 Utrecht University students (mean age: 23.6 ± 4.8) participated in the experiment and were rewarded with course credits or EUR 7,- per hour. 10 participants were pseudo-randomly assigned to the adaptation group and 10 to the control group. All participants

had normal or corrected to normal visual acuity and gave informed consent. Participants were debriefed about the adaptation procedure following completion of the experiment.

Materials

Stimuli were displayed on a 19" Nokia Multigraph 446Xpro CRT monitor with an effective display area of 36 x 27 cm, except for the first four participants in Experiment 2, for whom stimuli were displayed on a 21" Lacie Electron 22 Blue IV CRT monitor with an effective display area of 36 x 26 cm. Stimuli sizes and distances were computed to be identical on both monitors. Both monitors operated on a resolution of 1024x768 at a refresh rate of 120Hz. To avoid visual referencing cues from CRT afterglow, a 4-stop neutral density filter was attached to the screen. Monocular eye movements were recorded by an Eyelink1000 eye tracker (SR Research Ltd, Canada) on a Windows machine at a temporal resolution of 1000Hz and a spatial resolution of .01°. Participants were seated on an adjustable chair and placed their head on a chinrest. The screen was positioned 70cm from the front edge of the chinrest. To prevent visual guidance the experiment took place in a room with near-absolute darkness with the exception of stimuli and the Eyelink1000 illuminator glow. The experiment was designed with PyGaze (Dalmaijer et al., 2013), a python toolbox for eyetracking experiments based in this case on the PsychoPy and Pylink python libraries.

Procedure

Experiment 1 Two conditions were tested in each experiment session, a slow-paced and a fast-paced double step saccade task. All sessions began with verbal and on-screen instructions, an eye tracker calibration and practice trials for each condition, followed by a number of calibration trials (see below) for the fast condition to determine how fast participants could initiate saccades. Afterwards participants performed 4 blocks of 100 fast or slow trials each with breaks and eye tracker recalibration between blocks (fig. 1). Sessions lasted about 85 minutes.

All trials began with a blue fixation target (F; empty circle, diameter = .5°) appearing at a random location within the effective display area. In slow trials, after participants pressed the space bar and fixation was detected for 250ms, the fixation became a filled circle and

after a random delay of 800-1500ms a gray, filled target circle appeared. After 200ms a second gray target circle appeared. 200-800ms later the fixation dot disappeared, signaling participants to begin the saccade sequence. Target configurations are described in fig. 2B below. As soon as gaze was detected 3° outside of fixation the targets disappeared, leaving a blank screen at the end of the first saccade. The location of the second target therefore had to be remapped based on memory and extra-retinal cues.

In the fast condition (low pre-programming) the fixation circle became filled after participants pressed the space bar and after 500ms the second target appeared first on the screen. Then, after a delay of 300ms the first target appeared and the fixation circle disappeared, signaling participants to initiate saccades. If the saccade to the first target was not detected within a threshold period (see calibration procedure below) the screen was blanked, a warning beep sounded and the trial was repeated. Since the first target appeared after the second target this greatly reduced the time the visual system had available to pre-program the second saccade (e.g. Li & Andersen, 2001). Previous to the experimental trials, during threshold calibration, fast trial types were repeated until ten consecutive trials were executed without repeat. If a trial was repeated three times in a row, the threshold value was increased by 100ms, starting at 100ms up to a maximum of 500ms. Mean threshold was $360\text{ms} \pm 70\text{ms}$. During the experimental trials this served to force participants to begin programming the first saccade as soon as the first target appeared. During all trials, if gaze moved more than 2° from fixation before the go signal, the trial was terminated and repeated.

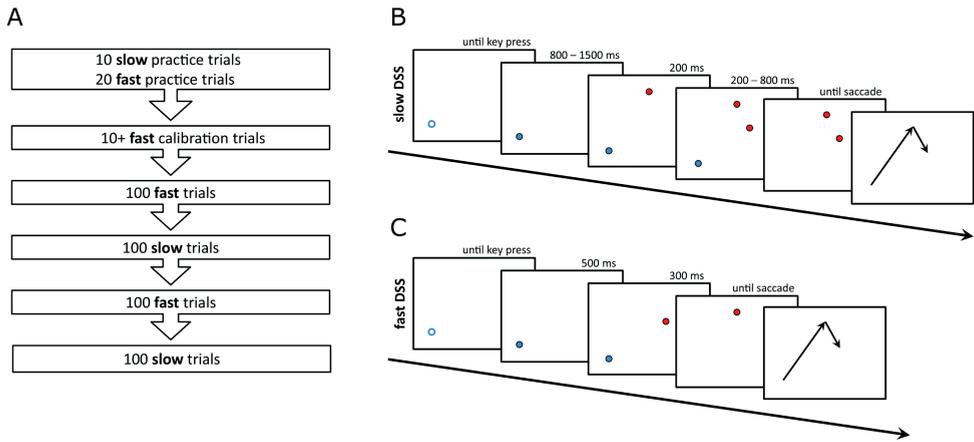


Figure 1: **A)** Experimental sequence Experiment 1. **B)** Slow paced DSS task. Colors and background changed for visual clarity. Arrows indicate expected eye movement path. **C)** Fast paced DSS task.

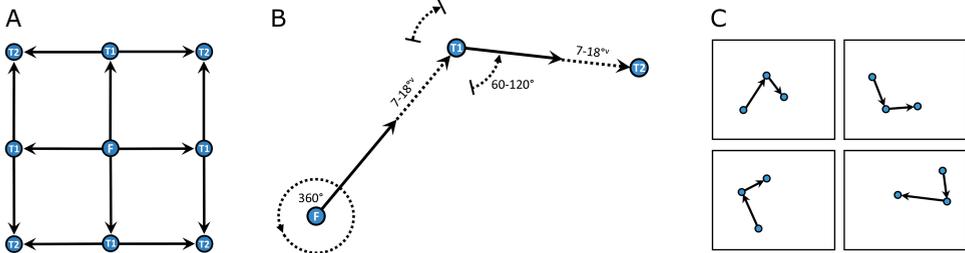


Figure 2: **A)** a typical array of possible stimulus configurations in previous DSS tasks. This creates only 4 possible endpoint locations. **B)** Possible target configuration in the present study. Fixation (F) could appear anywhere on the effective display area. The first target (T1) could appear anywhere around F but at a distance of least 7° and a maximum of 18° . The second target (T2) could appear at a distance of $7-18^\circ$ from T1, but with a polar angle between the first and second target vector of $60-120^\circ$. $^\circ$ = polar angle; $^\circ v$ = visual angle. **C)** Four example trial configurations.

Experiment 2 Participants received verbal and on-screen instructions and practiced the task with 15 adaptation and 15 DSS trials. 100 DSS trials measured baseline gain and were followed by 100 adaptation trials with a 25% back-step and 100 adaptation trials with a 35% back-step. Subsequently 50 cycles of 4 adaptation plus 2 DSS trials produced 300 test trials (fig.3). Sessions lasted about 75 minutes. The interleaved adaptation trials ensured that saccades remained adapted throughout the session. Participants were allowed to take short breaks (black screen) while remaining seated after 100, 200, 300 and 450 test trials with one recalibration after the first break. To ensure that adaptation effects were not caused by fatigue or learning (Golla et al., 2008), a control group performed the experiment without target back-step in the adaptation task.

To ensure maximum transfer of adaptation, in this experiment the DSS and adaptation tasks were designed to be as similar as possible in terms of stimuli and timing. All trials began with a green fixation target (F; filled circle, diameter = $.5^\circ$) appearing at a random location within the effective display area. After participants pressed the space bar to start the trial F turned red and after a random delay of 800-1600ms jumped to a new location. If gaze moved more than 2° from F after the dot had turned red, the trial was terminated and repeated.

During DSS trials a second target appeared after 100ms, signaling participants to initiate a saccade sequence to the targets in order of their appearance. When gaze was detected 3° outside the previously displayed F the targets disappeared so that T2 had to be targeted from memory after the first saccade. After a delay of 3200ms to allow for the execution of the saccades the next F appeared at the previous location of T2. Target configurations are described in fig. 2.

During adaptation trials participants saccaded to the first target as soon as it appeared with F simultaneously disappearing. When gaze was detected 3° outside the previously displayed fixation saccades were considered in progress and the target was displaced (stepped back) towards F. After a delay of 2000ms to allow for the execution of the saccade the next fixation target appeared at the location of the displaced target.

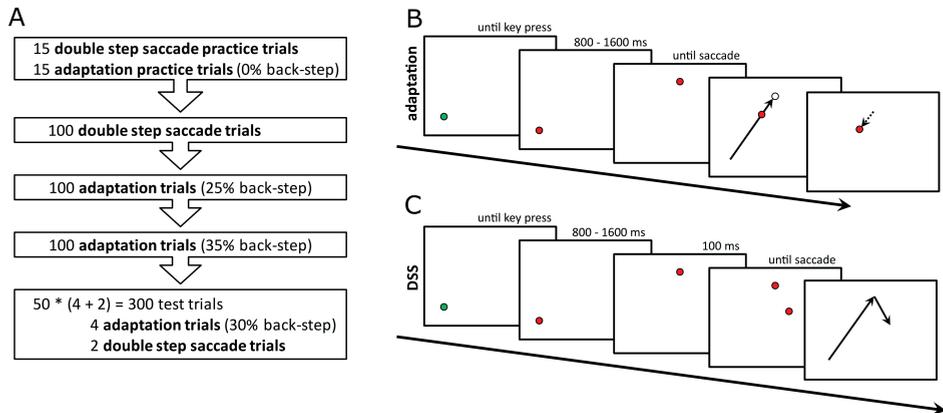


Figure 3: A) Experimental sequence Experiment 2. **B)** Adaptation trials. Colors and background changed for visual clarity. Arrows indicate expected eye movement path. **C)** DSS trials.

Data Analysis

Saccade detection was performed offline by SR Research software. Eye movements were considered saccades for velocities over $35^\circ/\text{s}$ or accelerations over $9500^\circ/\text{s}^2$. Further data analysis was performed in python. In each DSS trial a saccade to T1 (referred to here as S1) and a saccade to T2 (referred to here as S2) was defined from saccades with amplitude $> 4^\circ$. S1 was selected as the first saccade in the trial that followed T1 onset, started within a 2° radius around F and ended within 7° of T1. S2 was selected as the first saccade occurred after S1 if that ended within 7° of T2. An S1 made to T2 would lead to an S2 being made to T1, which would also invalidate the trial. If no two saccades in a trial satisfied these conditions, it was classified as invalid. These exclusion criteria resulted in a total loss of 13% of trials in Experiment 1 and 19% in Experiment 2. More conservative thresholds did not change the pattern of results reported but further reduced the dataset. The high proportion of invalid trials can be attributed to the difficulty of the task, in which stimulus locations were not predictable between trials. A subset of saccade configurations (around 10% of trials from different participants and conditions) were optically verified to ensure the validity of the

saccade selection method. 10 participants in Experiment 1 were replaced because more than 40% of their trials were invalid to ensure that unmotivated or unable participants did not confound the results. To avoid selection bias the analyses were repeated with all participants. This did not significantly alter the results.

Error compensation was computed as follows (also see fig.4): first saccade error can be described as a vector between the endpoint of the first saccade (E_1) and T_1 . If the visual system does not take this error into account but assumes to have landed on T_1 , it will program the second saccade as the vector between T_1 and T_2 . If this vector is executed from the actual gaze position E_1 , the endpoint of the second saccade will be placed at E_2' . Thus, the vector required for full compensation (e_1 ; the difference between E_2' and T_2) is identical to the first saccade error vector ($error_1$). This vector can be broken down into a projection onto the line between F and T_1 to produce a measure of ideal compensation in the direction of the first target vector (e_{1y}) and a perpendicular component (e_{1x}) to describe ideal compensation perpendicular to the first target vector. The same projections can be made for the vector from E_2' to the second saccade endpoint E_2 (e_2) to produce the actual amount of compensation in direction of the first target vector (e_{2y}) and perpendicular to it (e_{2x}). The regression slopes between e_{2y} and e_{1y} and between e_{2x} and e_{1x} reflect the average amount of error compensation in these two directions. The y components describe compensation in second saccades for error in in direction of the first target vector in first saccades (referred to here as amplitude error compensation). The x components describe compensation for error perpendicular to the direction of the first target vector in first saccades (referred to here as angular error compensation). A model for full compensation (where e_2 is identical to e_1 on every trial) is a perfect correlation between e_1 and e_2 if the second saccade does not introduce additional, unrelated error. Since there is always variance in saccades a true model for perfect compensation is the perfect correlation just described with variance added to e_2 (fig. 5 A). A model describing no compensation is a fully independent e_1 and e_2 (fig. 5 B). This provides a very direct and complete measure of error compensation in DSS as the entire error vector is taken into account. It is also unbiased by specific target vector amplitudes and angles. Linear regression was computed using the least-squares method, wherein slopes were optimized to minimize residuals in the vertical axis to give an estimate of the average compensation for endpoint error in the first saccade while disregarding unrelated variance of the second saccade. Average slopes between conditions were compared using

paired-samples two-tailed *t*-tests. S1 latencies were defined as the start of the first saccade minus offset of the fixation dot in the slow DSS condition and as the start of the first saccade minus onset of T1. Intersaccadic intervals were defined as start of the second saccade minus end of the first saccade in both conditions.

In Experiment 2 saccade gain ($S1_{GAIN}$ and $S2_{GAIN}$) was defined as the ratio between saccade amplitude and the distance between saccade starting point and target. Gain change after adaptation was analyzed by a 2-way MANOVA.

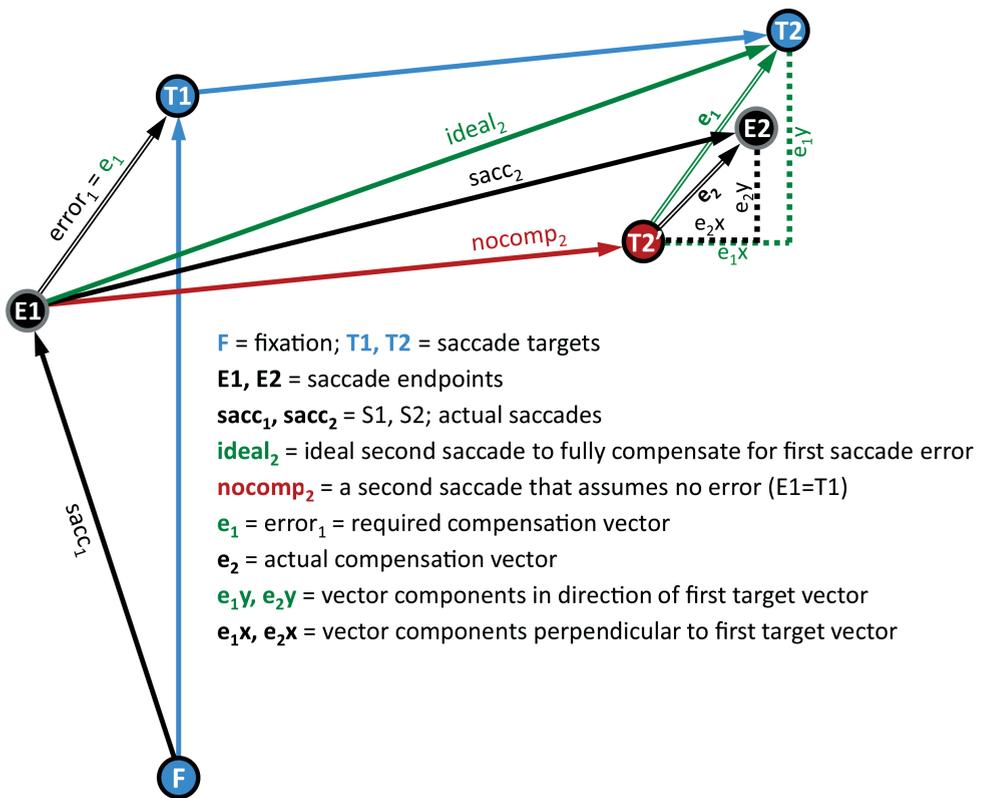


Figure 4: Geometric representation of the analysis method on an example trial.

RESULTS

Experiment 1 We investigated how saccades in a sequence correct for endpoint error of the previous saccade in a double step saccade task with random target vectors. To verify that the vector based analysis method is appropriate we produced two models using simulated data, one for no compensation (fig.5A) and one for full compensation (fig.5B). The no-compensation model was produced by regressing two independent random normal distributions, representing an uncorrelated e_1 and e_2 . Its regression slope is ~ 0 , indicating $\sim 0\%$ compensation. The full-compensation model was produced by plotting a random normal distribution (e_1) against itself ($e_2 = e_1$) with a random value from a separate random normal distribution added to the vertical coordinate of every point, representing an e_2 that perfectly follows e_1 (full compensation) but also introduces additional, non-systematic error. Its regression slope is ~ 1 , indicating $\sim 100\%$ compensation. Note that without additional (unrelated) error in the second saccade all points would lie exactly on the red and green line. In a case of perfect compensation every deviation of e_1 from 0 would be followed by an identical change in e_2 . Conversely, with no compensation and otherwise perfect S2 e_2 would always be 0. Since all saccades introduce error, we included random, vertical spread. The regression lines optimize vertical variance and indicate the average amount of compensation for error in S1.

In the fast DSS condition the first target appeared after the second target and saccades had to be initiated quickly after its onset. This would reduce the time the visual system had available to pre-program the second saccade along with the first saccade. The average regression slope from data of the ten participants was $.74 \pm .17$ (mean $R^2 = .37 \pm .14$) for amplitude error compensation, indicating $\sim 74\%$ compensation (blue lines in fig. 6A). Average slope for compensation of angular error (approximating angular error) was $.71 \pm .28$ (mean $R^2 = .16 \pm .14$), indicating 71% compensation (blue lines in fig. 6C). In the slow DSS condition participants were not forced to saccade quickly and thus could more efficiently pre-program both saccades from the retinal vector. In this condition the average regression slope was $.70 \pm .12$ (mean $R^2 = .35 \pm .07$) for amplitude error compensation, indicating $\sim 70\%$ compensation (red lines in fig. 6A). Average slope for angular error compensation was $.92 \pm .39$ (mean $R^2 = .27 \pm .18$), indicating 92% compensation (red lines in fig. 6C). These data verify

that saccade programming incorporates endpoint error of the previous saccade on a trial by trial basis, even when target vectors are randomly chosen for each trial.

Although there was a trend towards better compensation in the fast condition as seven out of ten participants showed this effect in amplitude error compensation (fig.6A&B) a paired-samples t-test revealed no significant evidence for a difference in compensation slopes between conditions for compensation of amplitude ($t(9) = 1.12$, $p = .29$) or angular error ($t(9) = -1.89$, $p = .09$).

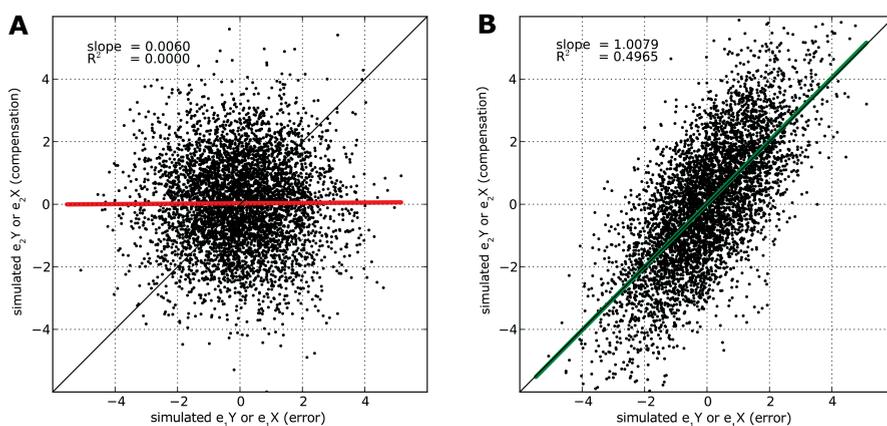


Figure 5: Simulated data from normal distributions with variance parameters of the observed data. **A)** Model for no compensation (independent e_1 and e_2). **B)** Model for full compensation (e_2 equals e_1 on average). Units are visual degrees.

Experiment 2 To understand whether systematic saccade amplitude errors are also taken into account in the programming of subsequent saccades we compared error compensation in DSS before and after global saccade adaptation.

Global saccade adaptation

We first examined the influence of the adaptation manipulation by comparing difference scores (change after adaptation) of S_{1GAIN} and S_{2GAIN} . A MANOVA identified significant gain changes on the group level (Wilks' Lambda = .266, $F(1,18) = 23.48$, $p < .001$, $\eta^2 = .734$) in

$\Delta S_{1\text{GAIN}}$ ($F(1,18) = 42.56, p < .0001, \eta^2 = .703$) and $\Delta S_{2\text{GAIN}}$ ($F(1,18) = 4.626, p = .045, \eta^2 = .204$). In other words, the gain change after adaptation differed significantly between adaptation and control group. The global adaptation procedure was successful and $S_{1\text{GAIN}}$ (.106 or 11.3% from baseline) and $S_{2\text{GAIN}}$ (.026 or 3.3% from baseline) in the DSS task were significantly reduced after the adaptation manipulation, but not in the control group. $S_{1\text{GAIN}}$ and $S_{2\text{GAIN}}$ tend to increase in the control group, which can be considered a learning effect whereby participants increase gain with practice. Adaptation is assumed to be working against this learning effect. This suggests that adaptation indeed transferred to S2. Fig.7 illustrates gain results as means of participant means before and after adaptation for both groups. Fig.8 A&B visualize gains before and after adaptation as a function of saccade direction. A timeline of adaptation is depicted in fig.8C. Amplitudes of saccades gradually reduced during adaptation trials and remained reduced throughout the post adaptation DSS trials. S1 latencies as measured from T1 onset (pre-adaptation: 318ms+120ms, post-adaptation: 300ms+116ms) confirmed that both targets were visible before S1 onset. These data demonstrate global saccadic adaptation for the trained saccade type (reactive) and suggests a moderate amount of transfer to memory guided saccades.

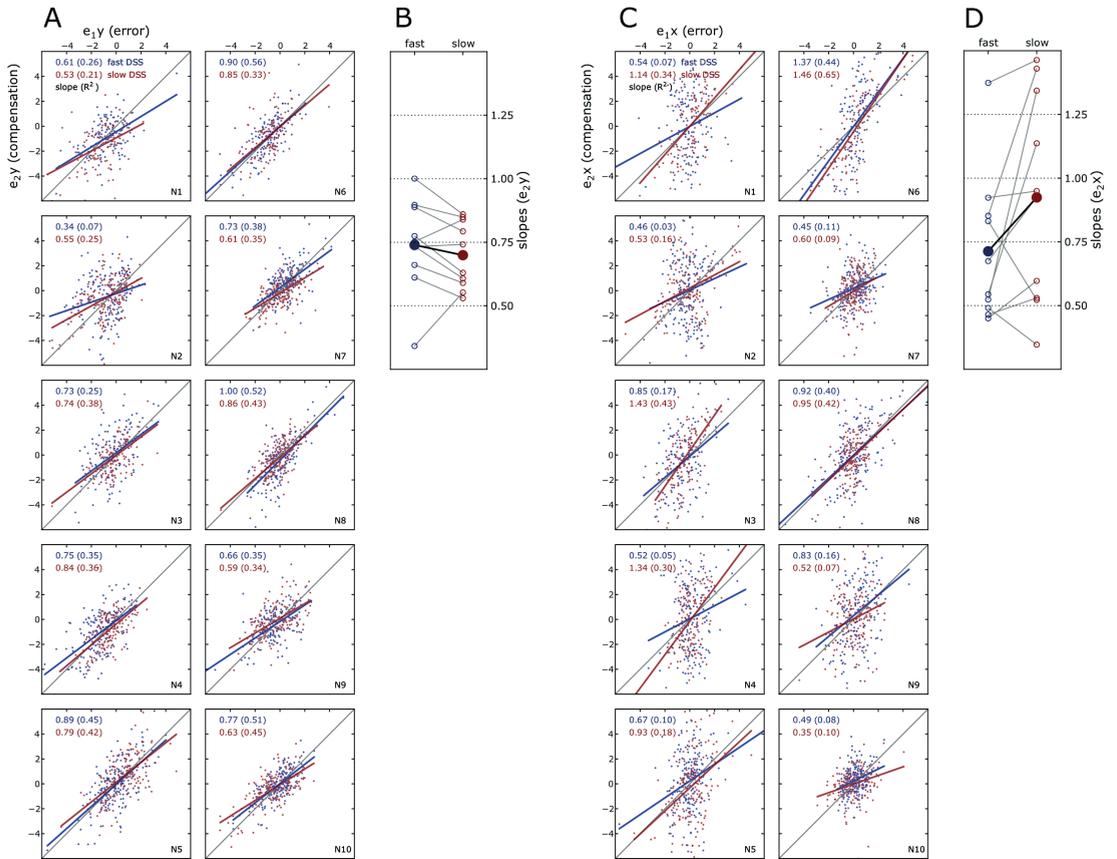


Figure 6: Individual regression slopes indicating average error compensation. The diagonal line represents full compensation. Units are visual degrees. **A)** Amplitude error compensation data per participant (for y projections). **B)** Replotted error compensation slopes (for y projections). Empty circles represent slopes of individual participants, filled circles represent group means. **C)** Angular error compensation per participant (for x projections). **D)** Replotted error compensation slopes (for x projections).

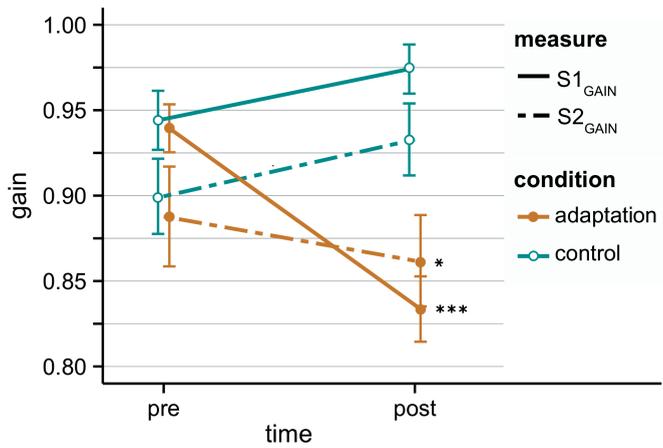


Figure 7: Means of participant means before and after adaptation in both groups. Bars represent standard error. $S1_{GAIN}$ and $S2_{GAIN}$ decreased in the adaptation group but not in the control group. * = sign. at .05 level. *** = sign. at .0001 level.

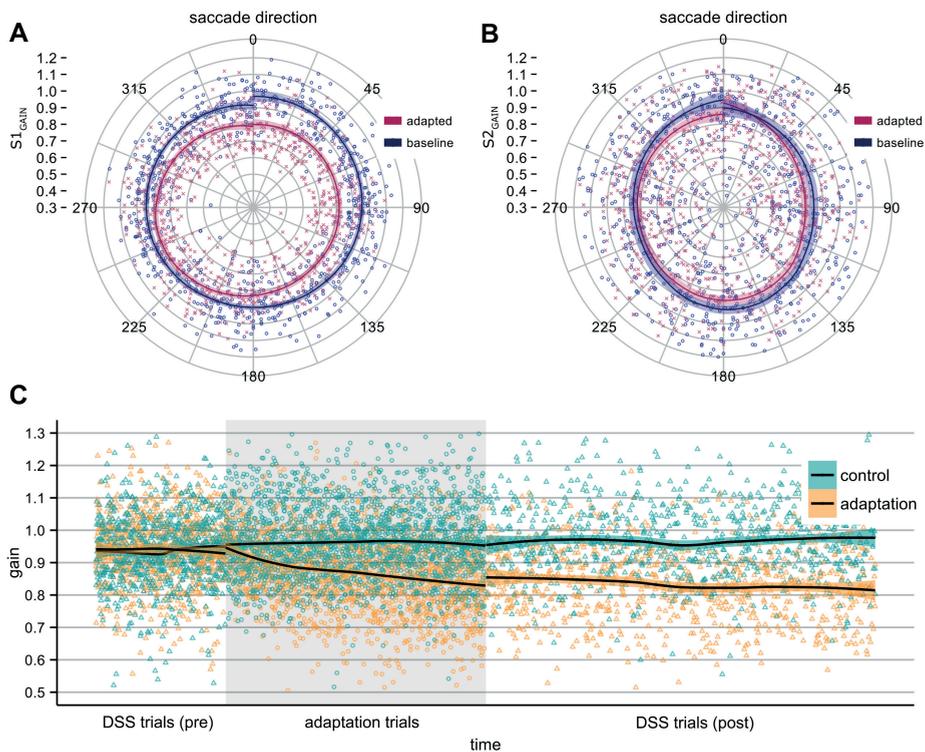


Figure 8: Gain as a function of saccade direction in the adaptation group before and after adaptation for **A)** S1 and **B)** S2 in a Local Polynomial Regression (LOESS) fit ($\alpha = .5$), folded into polar plots. Colored shadings represent 95% confidence intervals. **C)** Adaptation timeline. “DSS trials (post)” displays values for $S1_{GAIN}$ with the interleaved adaptation trials removed. Lines are LOESS fits ($\alpha = .8$). Colored areas represent 95% confidence intervals.

Error compensation in adapted DSS

Average regression slopes representing compensation for amplitude errors were not significantly different before (mean slope = $.58 \pm .38$, mean $R^2 = .19 \pm .17$) and after (mean slope = $.60 \pm .15$, mean $R^2 = .20 \pm .10$) adaptation (fig.9) in the adaptation group. Paired-samples two-tailed t -test: $t(9) = -.108$, $p = .917$. However, as can be seen in participant 4 (fig. 9A) an identical regression slope does not mean that the adapted vector is taken into account. After adaptation mean error (e_1y) was shifted to the left but mean compensation (e_2y) did not

shift downwards as would be expected if S2 were informed about the global reduction in S1 amplitudes. While in this participant compensation for S1 adaptation was not observed, fig.9B illustrates that most participants do take the adapted vector into account: an increase in error (shift to the left) is followed by an increase in compensation (shift downwards). Both the increase in error ($t(9)=6.04$, $p=.0002$) and the corresponding change in the compensation vector ($t(9)=3.36$, $p=.008$) were significant in repeated measures t-tests. This provides further evidence that during spatial remapping, extra-retinal signals from locations downstream of the site of adaptation that contain the adapted vectors enable the visual system to take this change in saccade metrics into account.

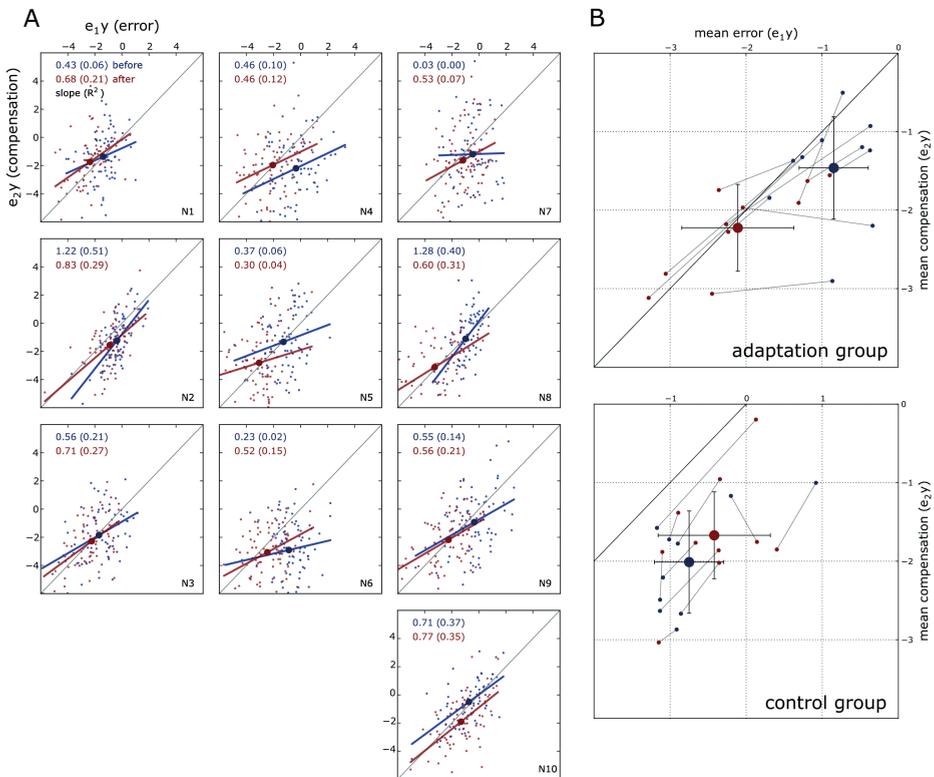


Figure 9: **A)** Individual regression slopes before and after adaptation in the adaptation group. **B)** Mean values for error and compensation vectors. Large circles indicate mean of means. Error bars represent one standard deviation.

DISCUSSION

In this study we investigated saccade programming in a sequence without visual guidance after the onset of the first saccade and the degree to which natural and induced saccade endpoint error is compensated in a subsequent saccade. Any eye movement displaces retinotopic target coordinates and requires the target vector to be recomputed using an extra-retinal signal of eye displacement and the memorized pre-saccadic target vector. Saccade error disallows full reliance on pre-programmed (predicted) post-saccadic target vectors and has to be taken into account to correctly saccade to the next target location. Regressing the difference between a hypothetical second saccade that ignores previous endpoint error and the actual saccade on the vector required for full error compensation produced a percentual estimate of the average error compensation per condition and participant.

We developed a double step saccade paradigm that displays targets at random locations to exclude saccade targeting based on long-term memory of the spatial layout of the stimuli. Previous studies investigating double step saccades used fixed target configurations (fig. 2A) that could easily have been memorized and saccades may have been corrected based on this memory instead of relying on CD. Joiner et al. (2010) reported very high error compensation for saccade pairs with identical amplitudes and a 90° angle separation between them. Angles of second saccades compensated around 100% of amplitude error in horizontal-vertical saccade pairs, 74% in vertical-horizontal pairs and 90% in oblique saccades. Collins (2010) used a similar spatial target layout and reported 74% compensation on average. In the first experiment of the present study we found comparable results of 74% amplitude error compensation in fast DSS and 70% in slow DSS, demonstrating that endpoint error correction in DSS sequences is not dependent on long-term spatiotopic memory of target configurations alone. These results provide further evidence for the existence of a CD signal and the reliance of the saccade system on this signal.

We aimed to induce a high degree of pre-programming by presenting the second target after the first, without limiting the time available to initiate the first saccade, in order to allow the visual system to plan both saccade vectors before the initiation of the first saccade. This condition was most comparable to traditional DSS paradigms with the exception that the two targets are typically presented

sequentially. Here their presentation time was partially overlapping to further increase pre-programming of the second saccade by providing clear visual vectors. In the fast DSS condition the second target was presented first and saccades had to be initiated immediately after the onset of the first target. A greater amount of compensation (steeper regression slopes) in this condition could have indicated that the visual system is aware of the decreased reliability of the pre-programmed second saccade and prioritizes information from a CD vector. However, differences in compensation between conditions were not significant. A limitation to these results is that the actual extent of pre-programming in the present DSS tasks cannot easily be quantified. A negative correlation between S1 latencies and ISI's has been suggested by Becker & Jürgens (1979) to indicate the extent of parallel processing prior to the onset of the first saccade. If participants had more time to pre-program both saccade vectors before initiating S1 then less time would be required to re-process S2 between saccades. For two reasons it is questionable whether this analysis can be applied to our results in Experiment 1. First, pre-programming would be expected in slow DSS. However, in this condition both targets were visible for a long time (200-800ms random fixation offset delay plus S1 latency). It can be assumed that any parallel processing of the two saccades would have been concluded prior to S1 onset and would therefore not affect ISI's. Second, Becker & Jürgens instructed participants to follow the target as soon as it jumped, causing saccades with longer reaction times to be targeted to the second target jump location directly. In contrast, we specifically instructed participants to saccade to the first and then to the second target. S1 latencies, intersaccadic intervals and their correlation within subjects are described in table 1. While the correlations are negative for nine of the ten subjects they are statistically significant in only 2. Further, paired-sample t-tests suggest that there was no difference in S1 latencies ($t(9) = .93, p = .38$). ISI's were even shorter in the fast condition ($t(9) = 2.97, p = .016$), although this could reflect the overall faster pace in this condition.

Future studies could investigate the precise relationship between parallel saccade processing and compensation by employing a continuous measure of pre-programming to address the question of how inputs from 'desired' and 'actual' saccade vectors are weighted to estimate eye displacement in the absence of visual feedback.

Table 1
Median first saccade latencies and intersaccadic intervals in Experiment 1.

ppt	Slow DSS			Fast DSS		
	S1 latency	ISI	r (p)	S1 latency	ISI	r (p)
1	343	224	0.02 (0.82)	303	173	-0.05 (0.57)
2	308	296	0.03 (0.66)	359	321	-0.22 (0.00)
3	266	322	0.22 (0.01)	252	242	-0.01 (0.89)
4	240	310	0.01 (0.95)	284	209	-0.05 (0.51)
5	284	261	0.07 (0.36)	273	256	-0.04 (0.59)
6	356	261	-0.12 (0.16)	403	198	-0.12 (0.20)
7	775	329	-0.16 (0.03)	298	222	-0.18 (0.01)
8	274	248	0.27 (0.00)	275	221	-0.10 (0.17)
9	319	245	0.05 (0.57)	264	273	0.11 (0.14)
10	423	323	0.13 (0.06)	420	228	-0.13 (0.09)

It is important to acknowledge the large variability between participants. For amplitude error compensation regression slopes ranged from .34 to 1.0. Variability in angle error compensation was even larger with slopes ranging from .34 to 1.42 in the slow DSS condition. In comparison, slopes ranged from .58 to 1.17 in the five participants in the study by Joiner et al. (2010). It is therefore not possible to converge on a global estimate of the amount of error compensation in the saccade system as this is highly variant from individual to individual. There appears to be within-subject consistency, however, as can be observed in the high correlation between compensation in the fast and slow DSS conditions of the present study, at least in amplitude error compensation (Pearson's $r = .784$, $p = .007$).

In Experiment 2 we investigated whether endpoint error caused by a systematic reduction of saccade amplitudes by adaptation is taken into account during error compensation. This is informative about possible neurophysiological sources of CD. If CD contains adapted vectors then saccade sequences should be executed correctly after adaptation. This is indeed what we found, in line with previous studies that reported veridical error compensation after adaptation of specific target vectors (e.g. Tanaka, 2003; Collins, 2010). Extra-retinal signals carrying adapted vectors must come from locations downstream from the site of adaptation. Converging evidence from lesioning studies (Goldberg et al., 1993; Barash et al., 1999; Takagi et al., 1998), spike recordings (Catz et al., 2005, 2008; Soetedjo and Fuchs, 2006) and neuroimaging (Desmurget et al., 1998; Desmurget et al., 2000) strongly points to the cerebellum as the most likely site of adaptation. The superior colliculus has been suggested to relay CD signals via the thalamus to frontal eye fields (Sommer & Wurtz, 2004) but lies upstream from the cerebellum, making it an unlikely source of extra-retinal signals containing adapted vectors. One study found

adaptation related activation in the SC (Takeichi et al., 2007), contradicting reports by Frens & van Opstal (1997). A more recent paper by Quessy et al. (2010) disagrees with the results by Takeichi et al. on methodological grounds and presents further evidence that amplitude encoding in SC is not altered by adaptation. Adaptation can transfer to saccades elicited by electric stimulation of the SC (Fitzgibbon & Goldberg, 1986; Edelman and Goldberg, 2002), further supporting the hypothesis that adaptation takes place downstream from the SC. Multiple pathways may contribute extra-retinal signals as lesioning the SC-MD-FEF pathway has only partially disrupted spatial remapping (Sommer & Wurtz, 2002, 2006). Indeed, several subcortical structures downstream from the site of adaptation in the cerebellum send ascending projections to the thalamus (for an overview see Tanaka, 2003). Our results contribute to the hypothesis that at least one source of extra-retinal signals lies at or downstream from the cerebellum.

Studies investigating transfer of adaptation from reactive to memory guided saccade types have shown mixed results. Deubel (1995, 1999) found no evidence of transfer whereas other studies such as Hopp & Fuchs (2002, 2010), Fujita et al. (2002) and Panouillères et al. (2014) reported adaptation transfer of 50-100% from reactive to memory guided saccades. In our study S1 were reactive saccades (made to a visible target at onset) while S2 were memory guided. In line with the latter studies we observed a moderate but significant gain reduction in S2 after the adaptation procedure as compared to the control group. This suggests that with global (parametric) adaptation transfer from reactive to memory guided saccade types may also be occurring.

In summary, we present, test and apply a novel, random-walk based DSS paradigm capable of examining saccades of all directions and amplitudes and suggest that endpoint error compensation in saccade sequences relies on trial-by-trial CD even for random target configurations.

REFERENCES

- Barash, S., Melikyan, A., Sivakov, A., Zhang, M., Glickstein & M., Their, P. (1999). Saccadic dysmetria and adaptation after lesions of the cerebellar cortex . *Journal of Neuroscience* , 19 , 10931 – 10939 .
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision research*, 19(9), 967-983.
- Catz, N., Dicke, P.W. & Thier, P. (2005). Cerebellar complex spike firing is suitable to induce as well as to stabilize motor learning. *Curr. Biol.* 15, 2179–2189.
- Catz, N., Dicke, P.W. & Thier, P. (2008). Cerebellar-dependent motor learning is based on pruning a Purkinje cell population response. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7309–7314.
- Colby, C. L., & Goldberg, M. E. (1999). Space and Attention in Parietal Cortex. *Annual Review of Neuroscience*, 22(1), 319–349. doi:10.1146/annurev.neuro.22.1.319
- Collins, T., Rolfs, M., Deubel, H., & Cavanagh, P. (2009). Post-saccadic location judgments reveal remapping of saccade targets to non-foveal locations. *Journal of Vision*, 9(5), 29.
- Collins, T. (2010). Extraretinal signal metrics in multiple-saccade sequences. *Journal of Vision*, 10(14), 7. doi:10.1167/10.14.7
- Dalmajer, E. S., Mathôt, S., & Van der Stigchel, S. (2013). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 1–9. doi:10.3758/s13428-013-0422-2
- Desmurget, M., Pélisson, D., Urquizar, C., Prablanc, C., Alexander, G.E. & Grafton, S.T. (1998). Functional anatomy of saccadic adaptation in humans. *Nat. Neurosci.* 1, 524–528.
- Desmurget, M., Pélisson, D., Grethe, J.S., Alexander, G.E., Urquizar, C., Prablanc, C. & Grafton, S.T. (2000). Functional adaptation of reactive saccades in humans: a PET study. *Exp. Brain Res.* 132, 243–259.
- Deubel, H. (1995). Separate adaptive mechanisms for the control of reactive and volitional saccadic eye movements. *Vision research*, 35(23), 3529-3540. doi:10.1016/0042-6989(95)00058-M
- Deubel, H. (1999). Separate mechanisms for the adaptive control of reactive, volitional, and memory-guided saccadic eye movements. *Attention and Performance XVII: Cognitive Regulation of Performance*.
- Ditterich, J., Eggert, T., & Straube, A. (1998). Fixation errors and timing in sequences of memory-guided saccades. *Behavioural brain research*, 95(2), 205-217. doi:10.1016/S0166-4328(97)00160-5
- Duhamel, Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040), 90–92. doi:10.1126/science.1553535
- Edelman, J. A., & Goldberg, M. E. (2002). Effect of short-term saccadic adaptation on saccades evoked by electrical stimulation in the primate superior colliculus. *Journal of Neurophysiology*, 87(4), 1915-1923.

- Frens, M. A., & Van Opstal, A. J. (1997). Monkey superior colliculus activity during short-term saccadic adaptation. *Brain research bulletin*, 43(5), 473-483.
- Fitzgibbon, E.J. & Goldberg, M.E. (1986). Cellular activity in the monkey superior colliculus during short term saccadic adaptation. *Soc. Neurosci. Abstr.* 325, 12.
- Fujita, M., Amagai, A., Minakawa, F., & Aoki, M. (2002). Selective and delay adaptation of human saccades. *Cognitive Brain Research*, 13(1), 41-52.
- Gabay, S., Pertzov, Y., Cohen, N., Avidan, G., & Henik, A. (2013). Remapping of the environment without corollary discharges: Evidence from scene-based IOR. *Journal of Vision*, 13(8), 22. doi:10.1167/13.8.22
- Goldberg, M.E., Musil, S.Y., Fitzgibbon, E.J., Smith, M.K. & Olson, C.R. (1993). The role of the cerebellum in the control of saccadic eye movements. In: Mano, N. (Ed.), *Cerebellum and Basal Ganglia in the Control of Movement*. Elsevier, Amsterdam, pp. 203–211.
- Golla, H., Tziridis, K., Haarmeier, T., Catz, N., Barash, S. & Thier, P. (2008). Reduced saccadic resilience and impaired saccadic adaptation due to cerebellar disease. *European Journal of Neuroscience*, 27(1), 132-144.
- Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983a). Corollary discharge provides accurate eye position information to the oculomotor system. *Science*, 221(4616), 1193–1195. doi:10.1126/science.6612334
- Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983b). Corollary discharge provides accurate eye position information to the oculomotor system. *Science*, 221(4616), 1193–1195. doi:10.1126/science.6612334
- Hallett, P. E., & Lightstone, A. D. (1976). Saccadic eye movements towards stimuli triggered by prior saccades. *Vision Research*, 16(1), 99–106. doi:10.1016/0042-6989(76)90083-3
- Herman, J. P., Blangero, A., Madelain, L., Khan, A., & Harwood, M. R. (2013). Saccade adaptation as a model of flexible and general motor learning. *Experimental Eye Research*, 114, 6–15. doi:10.1016/j.exer.2013.04.001
- Hopp, J. J., & Fuchs, A. F. (2002). Investigating the site of human saccadic adaptation with express and targeting saccades. *Experimental brain research*, 144(4), 538-548.
- Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in Neurobiology*, 72(1), 27–53. doi:10.1016/j.pneurobio.2003.12.002
- Hopp, J. J., & Fuchs, A. F. (2010). Identifying sites of saccade amplitude plasticity in humans: transfer of adaptation between different types of saccade. *Experimental brain research*, 202(1), 129-145.
- Iwamoto, Y., & Kaku, Y. (2010). Saccade adaptation as a model of learning in voluntary movements. *Experimental Brain Research*, 204(2), 145–162. doi:10.1007/s00221-010-2314-3
- Joiner, W. M., FitzGibbon, E. J., & Wurtz, R. H. (2010). Amplitudes and directions of individual saccades can be adjusted by corollary discharge. *Journal of Vision*, 10(2), 22. doi:10.1167/10.2.22
- Li, C. S. R., & Andersen, R. A. (2001). Inactivation of macaque lateral intraparietal area delays initiation of the second saccade predominantly from contralesional eye positions in a

- double-saccade task. *Experimental brain research*, 137(1), 45-57.
doi:10.1007/s002210000546
- McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics*, 2(8), 359–362. doi:10.3758/BF03210071
- Medendorp, W. P., Goltz, H. C., Vilis, T., & Crawford, J. D. (2003). Gaze-Centered Updating of Visual Space in Human Parietal Cortex. *The Journal of Neuroscience*, 23(15), 6209–6214.
- Noto, C., Watanabe, S. & Fuchs, A. (1999). Characteristics of simian adaptation fields produced by behavioral changes in saccadic gain and direction. *J. Neurophysiol.* 81, 2798–2813.
- Panouillères, M., Habchi, O., Gerardin, P., Salemmé, R., Urquizar, C., Farne, A., & Pélisson, D. (2012). A role for the parietal cortex in sensorimotor adaptation of saccades. *Cerebral Cortex*.
- Pélisson, D., Alahyane, N., Panouillères, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience & Biobehavioral Reviews*, 34(8), 1103-1120.
- Provis, J. M., Dubis, A. M., Maddess, T., & Carroll, J. (2013). Adaptation of the central retina for high acuity vision: cones, the fovea and the avascular zone. *Progress in Retinal and Eye Research*, 35, 63–81. doi:10.1016/j.preteyeres.2013.01.005
- Quessy, S., Quinet, J., & Freedman, E. G. (2010). The locus of motor activity in the superior colliculus of the rhesus monkey is unaltered during saccadic adaptation. *The Journal of Neuroscience*, 30(42), 14235-14244.
- Rolfs, M., Knapen, T., & Cavanagh, P. (2010). Global saccadic adaptation. *Vision Research*, 50(18), 1882–1890. doi:10.1016/j.visres.2010.06.010
- Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science*, 296(5572), 1480-1482.
- Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, 444(7117), 374–377. doi:10.1038/nature05279
- Soetedjo, R. & Fuchs, A.F., 2006. Complex spike activity of purkinje cells in the oculomotor vermis during behavioral adaptation of monkey saccades. *J. Neurosci.* 26, 7741–7755.
- Takagi, M., Zee, D.S. & Tamargo, R.J. (1998). Effects of lesions of the oculomotor vermis on eye movements in primate: saccades. *J. Neurophysiol.* 80, 1911–1931.
- Takeichi, N., Kaneko, C. R., & Fuchs, A. F. (2007). Activity changes in monkey superior colliculus during saccade adaptation. *Journal of neurophysiology*, 97(6), 4096-4107.
- Tanaka, M. (2003). Contribution of Signals Downstream From Adaptation to Saccade Programming. *Journal of Neurophysiology*, 90(3), 2080–2086.
doi:10.1152/jn.00207.2003

2.2 Remapping high-capacity, pre-attentive, fragile sensory memory.

ABSTRACT

Humans typically make several saccades per second. This provides a challenge for the visual system as locations are largely coded in retinotopic (eye-centered) coordinates. Spatial remapping, the updating of retinotopic location coordinates of items in visuospatial memory, is typically assumed to be limited to robust, capacity-limited and attention-demanding working memory (WM). Are pre-attentive, maskable, sensory memory representations (e.g. fragile memory, FM) also remapped? We directly compared trans-saccadic WM (tWM) and trans-saccadic FM (tFM) in a retro-cue change-detection paradigm. Participants memorized oriented rectangles, made a saccade and reported whether they saw a change in a subsequent display. On some trials a retro-cue indicated the to-be-tested item prior to probe onset. This allowed sensory memory items to be included in the memory capacity estimate. The observed retro-cue benefit demonstrates a tFM capacity considerably above tWM. This provides evidence that some, if not all sensory memory was remapped to spatiotopic (world-centered, task-relevant) coordinates. In a second experiment, we show backward masks to be effective in retinotopic as well as spatiotopic coordinates, demonstrating that FM was indeed remapped to world-centered coordinates. Together this provides conclusive evidence that trans-saccadic spatial remapping is not limited to higher-level WM processes but also occurs for sensory memory representations.

Published as:

Zerr, P., Gayet, S., Mulder, K., Pinto, Y., Sligte, I., & Van der Stigchel, S. (2017). Remapping high-capacity, pre-attentive, fragile sensory memory. *Scientific Reports*, 7(1), 1-10.

PZ, SG, YP, IS and SS designed the study concept. PZ programmed the experiment, collected experiment data, and wrote the manuscript. PZ and KM conducted the analyses. Critical revisions were provided by all co-authors.

INTRODUCTION

To only process visual information while it is available to the eyes would be a fatal disadvantage for dynamic agents in a dynamic world. A memory buffer allows us to access visual information after termination of its retinal input (visuospatial short-term memory; **VSTM**). Aristotle already mentioned the phenomenon of visible persistence in 3rd century B.C.¹, and the scientific study of VSTM goes back to at least Fechner² and Helmholtz³. It continues to captivate researchers to this day⁴. The discrepancy between the rich subjective experience of the world and the scientific evidence for the limits of perception is especially puzzling⁵.

The most commonly studied form of VSTM, stable **visuospatial working memory (WM)** is robust to visual masking⁶, can last tens of seconds but has a very limited capacity⁷ and requires dedicated attentional resources^{8,9}. The contents of VSTM are typically probed in change-detection experiments, which require observers to compare two arrays of visual stimuli that are presented in succession. However, memory traces that exist beyond the capacity of WM, so called **sensory memory**, are easily disrupted by new visual input. VSTM is thus composed of stable WM and unstable sensory memory. In traditional change-detection experiments, sensory memory traces are abolished by the presentation of the second array (i.e., the memory probe), such that only robust WM can be measured. The partial-report paradigm¹⁰ and the use of retro-cues^{11,12} have enabled researchers to investigate forms of sensory memory, such as **iconic memory**¹³ (**IM**): an unstable, pre-attentive, high-capacity memory trace that lasts a few hundred milliseconds. Prior to probe onset, a retro-cue indicates which items will be tested and directs internal attention (cognitive resources) to this item's memory representation. In this way an unstable memory trace can be stabilized, protected from being masked by the probe, and subsequently reported (for a review of the retro-cue paradigm see Souza and Oberauer¹⁴). This procedure dramatically increases capacity estimates because now both stable and unstable memory can be measured. Recently, a second form of sensory memory has been proposed to exist as an intermediate stage between IM and WM: **fragile memory (FM)**¹⁵. FM lasts several seconds and has a capacity that lies between IM and WM¹⁵. FM appears to be largely unimpaired by a withdrawal of attentional resources^{16,17}, which has led to the proposal that FM represents pre-attentive memory. At the very least, and in contrast to items in WM, FM items that can only be reported in a retro-cue paradigm have

received sufficiently fewer attentional resources to be masked by the visual interference of a probe array. While IM is masked by any visual stimulation such as a light flash or noise pattern, FM is only masked by visual input that resembles the memorized object⁶. This suggests that FM, unlike IM, is object-based: visual features (e.g. orientation) are bound to locations. FM has been demonstrated for both simple stimuli as well as more complex objects¹⁵. Whether FM really is a separate form of memory remains under debate¹⁸. We refer to FM as simply the increased capacity observed in the retro-cue paradigm and assume these additional items to have received little to no attentional resources. It is so far unknown whether FM survives eye movements, which forms the focus of the present study.

Humans typically make several fast eye movements (saccades) per second. This provides a challenge for the visual system as locations are largely coded in retinotopic (eye-centered) coordinates. To retain the actual location of an item, its memory representation needs to be updated with every saccade. Spatial remapping is the updating of retinotopic coordinates to task-relevant, spatiotopic (world-centered) coordinates¹⁹. It is so far unknown whether high-capacity sensory memory is remapped across saccades. The visual system could preserve resources by remapping only few items. Indeed, influential accounts of spatial remapping advocate that an item must have first received dedicated attentional resources before it can be remapped^{20–22}. This view proposes that only items in stable WM are maintained across eye movements. Reports from trans-saccadic VSTM experiments have suggested that with every saccade or other visual interruption we lose all information about the world except for three to four items that we actively attended to. Investigating trans-saccadic IM, Irwin²³ discovered that the number of letters they remembered scaled with set size when participants made no eye movements. After saccades, capacity remained at a ceiling of three to four items regardless of set size. Irwin concluded that while stable WM is maintained across eye movements, IM is not. Prime et al.²⁴ also reported a trans-saccadic capacity limit of three to four items for luminance and orientation, and noted that attended items were more likely to be remembered. This evidence seems to suggest that much of the initial representation of the visual world is lost after saccades.

The observed capacity limit of trans-saccadic memory, however, stands in stark contrast to our subjective experience of the world, which appears detailed, stable and continuous across saccades. There is evidence for the retention of highly detailed information

across saccades, such as precise location information. Peri-saccadic displacement of a saccade target typically goes unnoticed, even if the displacement is large²⁵. However, if the target was removed for a short time after the saccade has landed, even very small displacements could be noticed. Spatiotopically presented masks have also been found to reduce the capacity of sensory memory, suggesting that in addition to un-maskable WM, sensory memory representations also exist in spatiotopic coordinates^{26,23}. Germeys et al.²⁷ previously reported an increase in trans-saccadic memory capacity when using a retro-cue followed by a blank as compared to not using a blank. However, their paradigm did not allow to distinguish in which state of memory (stable or fragile) the items existed at the time of remapping. To investigate remapping of purely sensory memory the saccade target in the present study was placed outside of the memory array.

The present study investigated whether sensory memory is remapped to world-centered coordinates. For this purpose, we employed a retro-cue change-detection paradigm and required observers to make a saccade during the retention interval. We ask the question whether only items in robust WM are maintained across saccades or whether items in sensory memory are also maintained. If FM is remapped, then trans-saccadic memory capacity in the retro-cue condition (**tFM**) should be higher than trans-saccadic memory capacity in the post-cue condition (**tWM**). The capacity difference indicates the number of remapped sensory memory items (Figure 1). A small performance drop is expected in the saccade conditions relative to fixation conditions as the saccade target typically takes up about one item in memory²⁸. The results clearly support the hypothesis that sensory memory items have been remapped. In a second experiment, we presented backward masks to selectively disrupt memory representations in retinotopic or spatiotopic coordinates to investigate where and when FM representations exist across saccades. The purpose of the masks in this experiment was not to interfere with stimulus encoding, but to interfere with the memory signal of the stimulus, which is required for subsequent report. These masks can be thought of as interfering items, which replace fragile memory traces encoded in the locations in which they are presented. We observed both retinotopic and spatiotopic masks to be effective. This confirms that FM items do not only exist in retinotopic coordinates but are also remapped the spatiotopic coordinates.

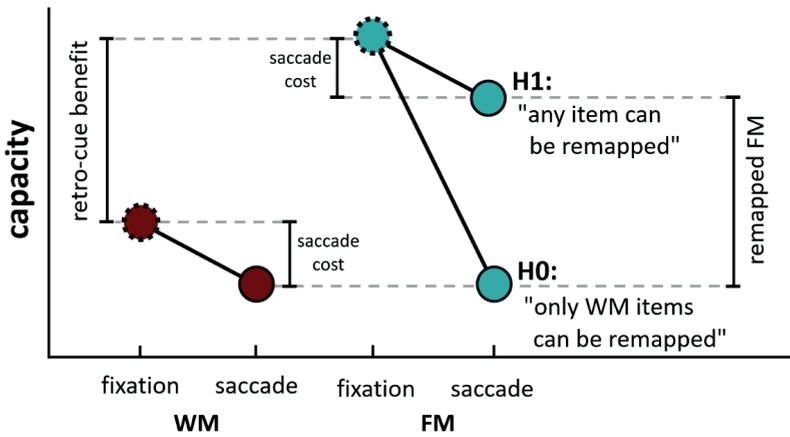


Figure 1: Hypotheses examined in Experiment 1. The retro-cue benefit demonstrates a higher memory capacity when using a retro-cue (i.e. FM). If this difference is also observed in the saccade condition, then this demonstrates trans-saccadic FM. Dashed circles indicate within-fixation capacity, circles with a solid outline indicate across-saccade capacity.

METHODS

Participants

All participants reported normal or corrected to normal visual acuity and gave informed consent. 8 Utrecht University students (aged 20-26, 6 female) participated in Experiment 1 and 21 students (aged 19-41, 15 female) participated in Experiment 2 for monetary reward. The study was approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University and has been carried out in accordance with the Declaration of Helsinki.

Materials

Stimuli were displayed in a dark room on an ASUS PG278q 27" LCD monitor with a display area of 60x34 cm (49.6x29.3 dva; degrees visual angle) and a resolution of 2560x1440 px at a refresh rate of 100 Hz and response time of 1ms. Monocular eye movements were recorded by an Eyelink1000 eye tracker (SR Research Ltd, Canada) at

a temporal resolution of 1000 Hz and a maximal spatial resolution of 0.01 dva. Participants were seated on an adjustable chair and placed their head on a chinrest 65 cm in front of the screen. The experiment was designed in Matlab 2015a and Psychtoolbox 3^{29,30}.

Stimuli and procedure

Experiment 1

Following verbal and on-screen instructions, participants completed the task with short breaks, approximately every 10-15 minutes. The eye tracker was re-calibrated at the beginning of the session and after each break. To ensure that the stimuli elicited no retinal afterimages we calibrated the gray value of the screen background to be perceptually isoluminant to the red (2.71 cd/m², x=.646, y=.338) stimuli for every participant. This was done by means of heterochromatic flicker photometry³¹.

Each trial began with three dots (0.22 dva) that always remained visible. The central fixation dot was flanked by two dark gray dots at a horizontal distance of 6.2 dva and remained blue until participants pressed the space bar to begin a trial, upon which the central dot turned red. The trial layout is illustrated in Figure 2. A memory array consisting of eight red rectangles in one of four orientations (0.25x0.9 dva, 2.5 dva from central fixation) was presented for 500 ms (Figure 2). On half the trials, after a blank interval of 100 ms (only the three fixation dots were visible), the red fixation dot jumped to the left or the right with the other two dots displayed in dark gray. On the other half of the trials the central dot remained red. Participants were required to keep their gaze at the red dot and to move their eyes immediately when it jumped to a new location. When their gaze deviated 2.5 dva from the red dot the trial was aborted and repeated at the end of the block. Participants subsequently pressed the space bar to go to the next trial. On half the trials, after 1000 ms to allow for both saccade execution and remapping, a valid retro-cue appeared as a red line (0.05 dva in width) from central fixation extending 1.44 dva in the direction of one of the memory items. After another blank of 1000 ms the probe array appeared again with either no change or with one of its items rotated 90°. This was always the cued item. Participants were instructed to indicate whether or not they noticed a change in the cued item by pressing the up/down arrow keys. After each trial participants received auditory feedback. The time delay between memory array

offset and retro-cue (the delay interval at which FM capacity is probed) was 1100 ms. On the other half of trials, the cue appeared together with the probe array (post-cue). This produced four conditions: post-cue with and without intervening saccade (i.e., probing WM), retro-cue with and without intervening saccade (i.e., probing FM). Crucially, the cue in the retro-cue condition appeared at the same time as the probe array in the post-cue condition, ensuring that WM and FM capacities could be adequately compared. Change/no-change as well as left/right saccade cue were balanced between conditions.

Experiment 2

See Figure 2 for a trial flow diagram. The second experiment was very similar to the first with three distinctions. First, masks were displayed after the saccade and before the retro-cue in order to selectively disrupt FM representations. These masks had the form of four randomly oriented bars, presented for 50 ms, followed by a blank for 50 ms and then followed by a second set of four randomly oriented bars for 50 ms. The masks overlapped the location of the future target in either a retinotopic location (around the current fixation dot) or a spatiotopic location (around the central dot). Presentation was either 50 ms (early) or 400 ms (late) after the saccade. The post-mask blank was presented for 650 ms (in the early condition) or 300 ms (in the late condition) after the mask to keep the delay interval identical between early and late masks. In two control conditions (tFM and tWM) masks were presented 400 ms after the saccade, in a spatiotopic location, containing no overlap with the future target. This ensured that a masking effect couldn't be explained by an increased working memory load. Second, five frames of colour noise were presented after the memory array, followed by a 300 ms blank to further eliminate the possibility that the effects observed in Experiment 1 could be explained by retinal afterimages. Third, the saccade cue remained on screen until gaze was detected 2.5 dva outside previous fixation. If no saccade was detected in this way after 500 ms the trial was aborted and repeated at the end of the block. This resulted in an average cue delay of 1300 ms + 217 ms (median saccade latency), almost half a second longer than in Experiment 1.

Data Analysis

Based on Cowan's k' we estimated memory capacity as:

$$(1) \quad k = [2 \times \textit{accuracy} - 1] \times \textit{memory set size}$$

It should be noted that the present study is agnostic with regard to working memory resources being allocated according to a discrete slot³² or continuous resource³³ model, a subject of ongoing debate. Cowan's k here can be seen as a continuous measure of capacity that is independent of set size. The term items in the present article therefore simply refers to a point on the Cowan's k scale. The essential metric is the difference in capacity between conditions within participants, rather than an absolute estimate of capacity limits in terms of number of discrete items.

Experiment 1

Sessions lasted 1.5 h in which participants performed 470 experimental trials on average, about 118 per condition. Participants practiced the memory task without the saccade condition for 1.5 h on the previous day.

Previous studies investigating FM have found that about one quarter of participants require excessive training or are unable to learn the task at all. It should be noted that this does not necessarily reflect individual differences in FM but more likely a failure to use the retro-cue, which requires participants to translate the exogenous retro-cue into an endogenous shift of attention. Since the focus of the present study was not within-fixation FM, two participants who did not supersede the pre-determined accuracy threshold of 85% in the within-fixation FM condition at the end of the training session were excluded from further participating in the experiment. This ensured reliable data for addressing the research question at hand: trans-saccadic FM.

Experiment 2

Sessions lasted 2 h, during which participants performed 530 experimental trials on average, about 88 per condition. Participants practiced the task without masks for 2 h on the previous day. In line with Experiment 1, 7 participants were excluded during the training

session if they did not reach 85% accuracy in the within-fixation FM condition after three blocks.

Statistical analysis

Bayesian analysis allows the use of a stopping rule³⁴. Data was collected until a preliminary analysis via a directional Bayesian t-test comparing the relevant conditions (tWM, tFM) reached $BF_{+0} = 6$. Subsequent analyses were performed in Stan³⁵ using R³⁶ and the R package brms³⁷. The model was a multilevel logistic regression using random intercepts to control for individual baseline differences with model equation

$$(2) \quad \text{logit}(\text{response}_{ij}) = b_0 + u_i + \sum_{con=1}^{n_{con}} b_{con} x_{con,j}$$

where response_{ij} is the predicted binary response (success) of participant i to trial j , b_0 is an intercept, which coincides with the mean log odds for the reference category, u_i is the random intercept for participant i , n_{con} is the total number of non-reference conditions, b_{con} is a difference parameter for condition con and $x_{con,j}$ is a dummy variable indicating whether trial j belongs to condition con .

Priors were set to be weakly informative for the first experiment. From previous research it is clear that participants will almost certainly do better than chance and will almost certainly not achieve perfect scores, which is reflected in the priors. The second experiment used the same random intercept model with different conditions. For this experiment, more informative priors were set based on the knowledge obtained in the first experiment. Priors are described in detail in the supplementary material.

We employed the Bayesian hypothesis testing framework using Bayes factors³⁸. For all hypothesis tests we first assessed the equality of two parameters, and subsequently tested for the direction of the effect^{39,40}. The equality hypothesis tests (H_0 : conditions are equal; H_1 : conditions are different) were computed in brms by the Savage-Dickey method⁴¹. The directed hypothesis tests are based on the ratio between the proportion of the posterior that is in agreement with a hypothesis and its complement. In other words, the ratio between the probability density mass of a difference between two conditions above and below

zero⁴²⁻⁴⁴. Posterior model probabilities were computed, for which prior odds for the hypothesis pairs under comparison were set to 1 to reflect equal weights of the hypotheses a priori. Analysis scripts and raw data are available via the Open Science Framework at <https://osf.io/ye9ya/>.

RESULTS

Experiment 1

In the first experiment, we demonstrate a retro-cue benefit irrespective of whether observers made a saccade. This provides strong evidence that FM items were remapped across saccades.

Higher capacity in the retro-cue condition

Accuracy scores for tFM were higher than tWM scores by 9.5 percentage points on average (0.48 ± 0.1 (mean \pm standard deviation) on the log-odds scale). This corresponds to an increase in memory capacity by $k = 1.5$ items. The Bayes Factor in favor of unequal tFM and tWM was 5530 with a posterior model probability of ≈ 1 . The directed hypothesis test indicated that tFM capacity is almost certainly larger than tWM capacity (posterior model probability ≈ 1).

Accuracy scores in the FM condition (within fixation) were 10.8 percentage points higher than WM scores (0.69 ± 0.12 on the log-odds scale). This corresponds to an increase in memory capacity by $k = 1.7$ items. The Bayes Factor in favor of unequal FM and WM was 26982 with a posterior model probability of ≈ 1 . The directed hypothesis test indicated that FM capacity is almost certainly larger than WM capacity (posterior model probability ≈ 1). Comparing tFM with FM capacity suggests that on average 1.5 of 1.7 sensory memory items were remapped.

Summary

We provide strong evidence that in addition to stable memory items some, if not all, fragile memory items have also been remapped. Individual data is shown in Figure 3 (left panel). The posterior probability density distributions are shown in Figure 3 (right panel). These distributions represent our knowledge about possible values of

the average proportion correct in each condition. Plots displaying the hypothesis tests graphically are provided in the supplementary material.

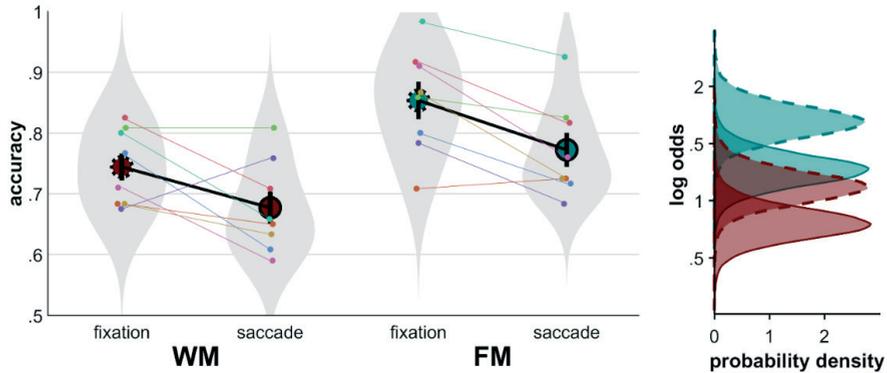


Figure 3: Results Experiment 1. (left) Individual proportions and means of proportions expressed as Cowan’s k . Dashed circles indicate within-fixation capacity, circles with a solid outline indicate across-saccade capacity. Error bars represent SE. Shaded gray areas are violin plots to visualize the shape of the distributions. (right) Posterior probability densities in log odds space. As can be seen, tFM capacity (cyan) is larger than tWM capacity (red).

Experiment 2

In the second experiment, we selectively masked locations to disrupt FM representations in order to assess where in space and time FM exists across saccades. The informative priors employed in this section are discussed in the supplementary materials. The results show that masks at the retinotopic location as well as at the spatiotopic location reduced the retro-cue benefit. This confirms that FM items have been remapped to task-relevant, world-centered coordinates.

Higher capacity in the retro-cue condition

First, we replicated the main finding of Experiment 1. On average, tFM accuracy scores were 10.1 percentage points higher than tWM scores (0.51 ± 0.07 on the log-odds scale). The Bayes Factor in favor of unequal tFM and tWM was 28408, with a posterior model probability of ≈ 1 . The directed hypothesis test again indicated that tFM capacity

is almost certainly larger than tWM capacity with a Bayes Factor approaching infinity and posterior model probability ≈ 1 .

No difference between early and late masks

For retinotopic masks, the Bayes Factor was 4.3 in favor of no difference between early and late mask presentation, with a posterior model probability of 0.811. Retinotopic early masks were 1.5 percentage points more effective than retinotopic late masks. For spatiotopic masks, the Bayes Factor was 7.5 in favor of no difference between early and late mask presentation, with a posterior model probability of 0.883. Spatiotopic late masks were 0.3 percentage points more effective than spatiotopic early masks. Following this evidence for no difference, early and late mask data were pooled for retinotopic and spatiotopic mask conditions respectively.

Retinotopic and spatiotopic masks are effective

On average, accuracy scores in the retinotopic mask condition were 4.8 percentage points lower than tFM scores ($0.25 \pm .06$ on the log-odds scale). The Bayes Factor in favor of unequal tFM and retinotopic mask conditions was 600, with a posterior model probability of 0.998. The directed hypothesis test indicated that retinotopic masks almost certainly were effective with a Bayes Factor approaching infinity and posterior model probability ≈ 1 .

On average, accuracy scores in the spatiotopic mask condition were 3.6 percentage points lower than tFM scores (0.19 ± 0.06 on the log-odds scale). The Bayes Factor in favor of unequal tFM and spatiotopic mask conditions was 16.84, with a posterior model probability of 0.944. The directed hypothesis test indicated that spatiotopic masks almost certainly were effective with a Bayes Factor of 1999 and a posterior model probability ≈ 1 .

Summary

We confirm that (1) tFM capacity is larger than tWM capacity and (2) FM is remapped to spatiotopic coordinates. Individual data is shown in Figure 4 (left panel). The posterior probability densities, which represent our knowledge about possible values for the parameters and display the effects found graphically, are shown in Figure 4 (right

panel). Plots displaying the hypothesis tests graphically are provided in the supplementary material.

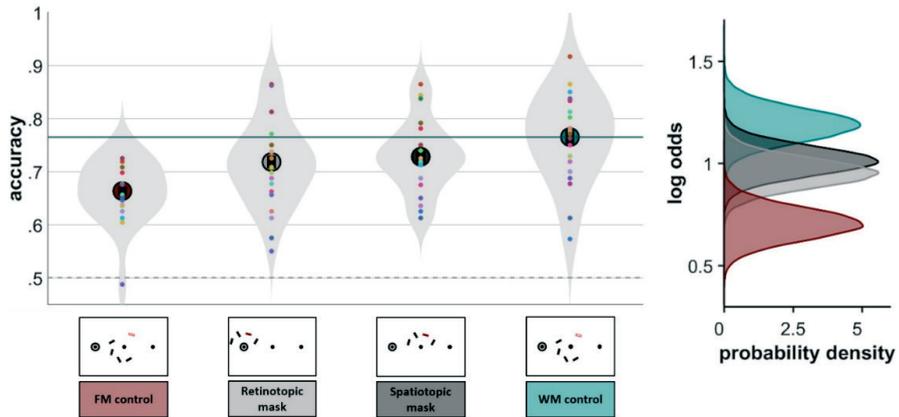


Figure 4: Results Experiment 2. (left) Proportions correct per participant and means of proportions correct per condition. Dashed circles indicate within-fixation capacity, circles with a solid outline indicate across-saccade capacity. Error bars represent SE. Shaded gray areas are violin plots to visualize the shape of the distributions. The mask type used is visualized below the violin plots of each condition. The transparent bar indicates the future target location. (right) Posterior probability densities in log odds space, which provide strong support that both spatiotopic and retinotopic masks interfered with memory representations (difference between cyan and gray distributions), as well as a replication of Experiment 1 (difference between cyan and red distributions).

DISCUSSION

We investigated trans-saccadic visuospatial short-term memory (tVSTM) in a retro-cue change-detection paradigm and observed remapping of both robust and fragile memory traces. In addition to items that received dedicated attentional resources (robust WM), some items in unstable, pre-attentive FM were also remapped. Our results challenge the strongly held beliefs that (1) remapping always requires attentional resources, and (2) that tVSTM contains only WM items and thus is synonymous with tWM.

The capacity of trans-saccadic VSTM

Previously, tVSTM capacity has been estimated to underlie the same limitations as WM: three to four items^{23,24}. According to this account, items that did not receive sufficient attentional resources prior to the saccade are lost after the saccade. We argue that studies investigating tVSTM have suffered from similar limitations as early studies of VSTM: there is more information in memory than can be reported in common paradigms⁵. Partial-report enabled the study of other forms of VSTM, such as sensory memory. Due to the fast decay of iconic memory it was only after the discovery of FM¹⁵, which has a lifetime of several seconds, that a paradigm could be developed to adequately estimate trans-saccadic sensory memory. Recently a study by Germeys et al.²⁷ reported a post-saccadic retro-cue benefit. However, in their paradigm the state of memory items at the time of remapping is unclear because the saccade target was always one of the memory items. To investigate remapping of purely sensory memory the saccade target in the present study was placed outside of the memory array.

We observed a trans-saccadic sensory memory capacity well above that of tWM. This is a clear indication that some, if not all, FM items have been remapped. Due to individual differences in WM capacity it is crucial to compare capacities within-subject and within-task, that is, WM after a saccade with FM after a saccade, rather than to establish global capacity estimates such as the “magic number four”. With an appropriate baseline condition (tWM) any capacity increase observed for tFM must reflect items that were stored as FM but remained accessible in task-relevant coordinates after the saccade (see also Figure 1). If only stable items survived saccades, then no retro-cue

benefit should have been observed because there would not have been any fragile memory traces left for the retro-cue to rescue.

In our view, previous studies^{23,24} have failed to observe remapping of sensory memory because the employed paradigms were not sensitive enough to pick up these fragile memory traces across saccades. Irwin²³ used a post-cue delayed-recall procedure and showed stimuli only once, requiring participants to begin recalling the test item from the set of memorized stimuli immediately after the post-cue. The intrusion-errors Irwin reported in Table 4 are much more pronounced for the saccade condition, which indicates difficulties to localize the test item in memory and explains the reduced performance after saccades. While Prime et al.²⁴ used a change-detection task, they did not include a retro-cue. Sensory memory can't be directly reported but must first be transferred to stable WM. The retro-cue change-detection task employed in the present study enabled participants to first stabilize a remapped fragile memory trace and to then compare it with the visible test item at the same spatiotopic location.

Spatiotopic and retinotopic masking

It could have been possible that participants in Experiment 1 were able to use the retro-cue without remapping the memory items. To ensure that FM was indeed remapped to spatiotopic coordinates we presented masks after the saccade and before the retro-cue in Experiment 2 and found both retinotopic and spatiotopic masks to be effective, in line with previous studies^{23,26}. Pinto et al.⁶ previously demonstrated that object-specific masks interfere with FM but not with WM, providing a tool for assessing the spatial properties of FM, namely, whether it is coded in a retinotopic or spatiotopic reference frame. Crucially, in the present study the masking effect was calculated relative to a condition with the same masking stimuli presented at locations that did not overlap with the future target. Since the masking stimuli were presented in all conditions and the only difference being their precise spatial location the masking effect cannot be explained by additional working memory load, which was a possibility in earlier studies^{23,26} or an interference with the comparison process in the change-detection task. The effectiveness of spatiotopic masks, together with the observed retro-cue benefit after the saccade, provides conclusive evidence that FM does not only survive eye movements but is indeed remapped to world-centered, task-relevant coordinates. Surprisingly, there was no difference between presenting masks early or late after

the saccade. We expected primarily early retinotopic masks and late spatiotopic masks to be effective as presumably memory representations first exist in retinotopic coordinates and are then remapped to spatiotopic coordinates. However, similar patterns of results have previously been reported. Golomb, Chun and Mazer⁴⁵ observed a lingering retinotopic trace of attentional pointers after a saccade, although for a much shorter duration than what was observed in the present study. In contrast to Golomb et al., who reported lingering retinotopic attentional activation, we observed an apparent lingering retinotopic memory representation. Conversely, while build-up of spatiotopic representations is often thought to be a slow process⁴⁶, immediate build-up has also recently been observed⁴⁷.

Pre-attentive spatial remapping and the functional role of sensory memory

Traditionally, trans-saccadic memory was believed to be comprised of WM items only since the spatial remapping process is thought to strongly depend on attention. Attention here refers to the allocation of a limited cognitive resource to a small set of objects. Items in WM have been prioritized by attentional mechanisms, while items in FM may have received no attention at all or sufficiently fewer resources to remain in a sensory memory state, which is easily masked by visual interference^{16,17}.

Especially the failure to detect a role for sensory memory in trans-saccadic perception has led this memory form to be dismissed as an artefact of visual processing⁴⁸. By demonstrating remapping of sensory memory items that have not received dedicated cognitive resources, we confirm that sensory memory is not merely an artefact, but an object-based form of memory that is maintained and updated across eye movements. This suggests potential functional roles for sensory memory. One possibility is that trans-saccadic perception relies on a highly detailed pre-categorical representation of the world. It is conceivable that remapping an entire scene across eye movements provides a reference frame for trans-saccadic integration and a computationally efficient way to localize the most relevant objects in memory (WM) within this frame. In that case, sensory memory would support trans-saccadic visual stability, one of the great puzzles of visual neuroscience⁴⁹.

An important question concerns whether FM items (defined as items that can only be reported in a retro-cue paradigm) depend on

qualitatively different resources (i.e., a different memory store) or on quantitatively different resources (i.e., receiving less attention) than WM items. We interpret our results within a framework that assumes the former possibility (FM and WM rely on a separate cognitive resource). Two aspects would be necessary to substantiate this framework: **(1)** a differential effect of withdrawal of attention on WM and FM. **(2)** Separate neural substrates of the two memory forms.

(1) Vandenbroucke, Sligte & Lamme¹⁶ found that reducing available attentional resources affected WM and FM very differently. They conducted three experiments to withdraw attention during a change detection task. Temporal uncertainty, a parallel n-back task and an attentional blink paradigm all demonstrated differential effects of attention withdrawal on WM and FM. Whereas WM capacity was reduced by withdrawing attention, FM capacity was reduced only minimally. The authors (or we) do not argue that FM operates independently of neural resources, as that would preclude any brain activity, but that FM operates independently of the resource that WM depends on (i.e. “attention”). Crucially, qualitative differences in capacity reduction between WM and FM were observed: a 2-back task impacted WM stronger than a 1-back task. This was not the case for FM. In other words, the accuracy reduction did not scale with the level of attentional diversion. The attentional blink impacted WM depending on lag position. FM was impacted identically at any lag position. In other words, the impact on FM was likely driven by the dual-task rather than the attentional blink as such. The observed numerical and qualitative differences make a strong case for WM and FM relying on separate cognitive resources and FM being a pre-attentive form of VSTM. Vandenbroucke et al.¹⁶ argued that FM represents the initial storage capacity of VSTM while WM depends on attentional boosting of some of these items for cognitive manipulation and report. This conclusion is in line with the initial study by Landman et al.¹¹, who demonstrated that even after an item had initially been attentionally selected by a retro-cue, a benefit from a second retro-cue, selecting a different item from the memory set, could still be retrieved. Furthermore, Pinto et al.¹⁷ observed a greater performance reduction in WM than FM when spatial attention was withdrawn from the future target during encoding. They presented 80% valid pre-cues before the memory array and examined the invalidly cued trials to measure the effect of attentional resource withdrawal. They concluded from the pattern of results (constant performance reduction in both paradigms) that the WM component was affected in both cases, indicating that WM and FM are separate rather than redundant memory stores.

(2) Sligte, Wokke, Tesselaar, Scholte and Lamme⁵⁰ discovered that TMS pulses to the right DLPFC strongly affected measures of WM but not FM. Sligte et al. were asking the question whether FM is a weak form of WM (items that received less attention); or whether FM is independent of the attentional boosting that WM relies on. Their results support the hypothesis that FM does not depend on the wide-spread network activity associated with the attentional boosting observed in WM, but relies at least partially on a different neurological architecture.

Taken together these results strongly suggests that FM does not depend on attention in the way WM does and that FM items received little to no attention during encoding. Irrespective of whether or not the FM interpretation of the retro-cue benefit effect holds, the present results at the very least demonstrate that items that exist in a lower priority state (fewer cognitive resources, less attentional prioritization, high maskability) are remapped in addition to those items that were prioritized and received most resources.

Conclusion

In summary, the present study provides compelling evidence for spatial remapping of items in pre-attentive, high-capacity, fragile sensory memory across eye movements. This contradicts previous studies that confine remapping to robust, capacity-limited WM and has implications for the role of attention in spatial remapping.

REFERENCES

1. Allen, F. The persistence of vision. *Am. J. Physiol. Opt.* (1926).
2. Fechner, G. T. Ueber die subjectiven Nachbilder und Nebenbilder. *Ann. Phys.* **126** (6), 193–221 (1840).
3. Helmholtz, H. V. Concerning the perceptions in general. *Treatise Physiol. Opt.* **3**, (1866).
4. Luck, S. J. & Vogel, E. K. Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* **17**, 391–400 (2013).
5. Cohen, M. A., Dennett, D. C. & Kanwisher, N. What is the Bandwidth of Perceptual Experience? *Trends Cogn. Sci.* **20**, 324–335 (2016).
6. Pinto, Y., Sligte, I. G., Shapiro, K. L. & Lamme, V. a. Fragile visual short-term memory is an object-based and location-specific store. *Psychon. Bull. Rev.* **20**, 732–739 (2013).

7. Cowan, N. The magical number 4 in short term memory. A reconsideration of storage capacity. *Behav. Brain Sci.* **24**, 87–186 (2001).
8. LaBar, K. S., Gitelman, D. R., Parrish, T. B. & Mesulam, M. Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *Neuroimage* **10**, 695–704 (1999).
9. Awh, E., Jonides, J. & Awh, E. Overlapping mechanisms of attention and spatial working memory. **5**, 119–126 (2001).
10. Sperling, G. The information available in brief visual presentations. *Psychol. Monogr.* **74**, 1–29 (1960).
11. Landman, R., Spekreijse, H. & Lamme, V. A. F. Large capacity storage of integrated objects before change blindness. *Vision Res.* **43**, 149–164 (2003).
12. Griffin, I. C. & Nobre, A. C. Orienting attention to locations in internal representations. *J. Cogn. Neurosci.* **15**, 1176–1194 (2003).
13. Neisser, U. *Cognitive Psychology*. (Appleton-Century-Crofts., 1967).
14. Souza, A. S. & Oberauer, K. In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, Psychophys.* 1–22 (2016). doi:10.3758/s13414-016-1108-5
15. Sligte, I. G., Scholte, H. S. & Lamme, V. A. F. Are there multiple visual short-term memory stores? *PLoS One* **3**, 2–10 (2008).
16. Vandembroucke, A. R. E., Sligte, I. G. & Lamme, V. A. F. Manipulations of attention dissociate fragile visual short-term memory from visual working memory. *Neuropsychologia* **49**, 1559–1568 (2011).
17. Pinto, Y. *et al.* Conscious Visual Memory With Minimal Attention. **146**, 214–226 (2016).
18. Matsukura, M. & Hollingworth, A. Does visual short-term memory have a high-capacity stage? *Psychon. Bull. Rev.* **18**, 1098–1104 (2011).
19. Bays, P. M. & Husain, M. Spatial remapping of the visual world across saccades. *Neuroreport* **18**, 1207–1213 (2007).
20. Irwin, D. E. & Gordon, R. D. Eye Movements, Attention and Trans-saccadic Memory. *Vis. cogn.* **5**, 127–155 (1998).
21. Cavanagh, P., Hunt, A. R., Afraz, A. & Rolfs, M. Visual stability based on remapping of attention pointers. *Trends Cogn. Sci.* **14**, 147–53 (2010).
22. Jonikaitis, D., Szinte, M., Rolfs, M. & Cavanagh, P. Allocation of attention across saccades. *J. Neurophysiol.* **109**, 1425–34 (2013).
23. Irwin, D. E. Memory for position and identity across eye movements. *Journal Of Experimental Psychology-Learning Memory And Cognition* **18**, 307–317 (1992).
24. Prime, S. L., Tsotsos, L., Keith, G. P. & Crawford, J. D. Visual memory capacity in transsaccadic integration. *Exp. Brain Res.* **180**, 609–28 (2007).
25. Deubel, H., Schneider, W. X. & Bridgeman, B. Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Res.* **36**, 985–996 (1996).
26. McRae, K., Butler, B. E. & Popiel, S. J. Spatiotopic and retinotopic components of iconic memory. *Psychol. Res.* **49**, 221–227 (1987).
27. Germeys, F., De Graef, P., Van Eccelpoel, C. & Verfaillie, K. The visual analog: evidence for a preattentive representation across saccades. *J. Vis.* **10**, 9 (2010).
28. Schut, M. J., Van der Stoep, N., Postma, A. & Van der Stigchel, S. The cost of making

- an eye movement: A direct link between visual working memory and saccade execution. *J. Vis.* **17**, 15 (2017).
29. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
 30. Kleiner, M. *et al.* What's new in Psychtoolbox-3? *Perception* **36**, S14 (2007).
 31. Kaiser, P. K. & Comerford, J. P. Flicker photometry of equally bright lights. *Vision Res.* **15**, 1399–1402 (1975).
 32. Zhang, W. & Luck, S. J. Discrete fixed-resolution representations in visual working memory. *Nature* **453**, 233–U13 (2008).
 33. Bays, P. M. & Husain, M. Dynamic shifts of limited working memory resources in human vision. *Science* **321**, 851–4 (2008).
 34. Rouder, J. N. Optional stopping : No problem for Bayesians. 301–308 (2014). doi:10.3758/s13423-014-0595-4
 35. Carpenter, B. *et al.* Journal of Statistical Software Stan : A Probabilistic Programming Language. *J. Stat. Softw.* **VV**, (2016).
 36. Ihaka, R. & Gentleman, R. Interface Foundation of America R: A Language for Data Analysis and Graphics R: A Language for Data Analysis and Graphics. *Source J. Comput. Graph. Stat.* **5**, 299–314 (1996).
 37. Burkner, P.-C. brms : An R Package for Bayesian Generalized Linear Mixed Models using Stan. (2015).
 38. Kass, R. E. & Raftery, A. E. Bayes Factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
 39. Marsman, M. & Wagenmakers, E. Three Insights from a Bayesian Interpretation of the One-Sided P Value. 1–11 (2016). doi:10.1177/0013164416669201
 40. Johnson, D. The insignificance of statistical significance testing. *J. Wildl. Manage.* **63**, 763–772 (1999).
 41. Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H. & Grasman, R. Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cogn. Psychol.* **60**, 158–189 (2010).
 42. Klugkist, I., Laudy, O. & Hoijtink, H. Inequality constrained analysis of variance: a Bayesian approach. *Psychol. Methods* **10**, 477–93 (2005).
 43. Klugkist, I., Kato, B. & Hoijtink, H. Bayesian model selection using encompassing priors. *Stat. Neerl.* **59**, 57–69 (2005).
 44. Wetzels, R., Grasman, R. P. & Wagenmakers, E. J. An encompassing prior generalization of the Savage-Dickey density ratio. *Comput. Stat. Data Anal.* **54**, 2094–2102 (2010).
 45. Golomb, J. D., Chun, M. M. & Mazer, J. A. The Native Coordinate System of Spatial Attention Is Retinotopic. *J. Neurosci.* **28**, 10654–62 (2008).
 46. Zimmermann, E., Morrone, M. C., Fink, G. R. & Burr, D. Spatiotopic neural representations develop slowly across saccades. *Curr. Biol.* **23**, R193–R194 (2013).
 47. Fabius, J. H., Fracasso, A. & Van der Stigchel, S. Spatiotopic updating facilitates perception immediately after saccades. *Sci. Rep.* **6**, 34488 (2016).
 48. Haber, R. N. The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behav. Brain Sci.* **6**, 1 (1983).
 49. Melcher, D. & Colby, C. L. Trans-saccadic perception. *Trends Cogn. Sci.* **12**, 466–73 (2008).

RESEARCH ARTICLES

50. Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Steven Scholte, H. & Lamme, V. A. F. Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia* **49**, 1578–1588 (2011).

2.3 The development of retro-cue benefits with extensive practice: implications for capacity estimation and attentional states in visual working memory

ABSTRACT

Accessing the contents of visual short-term memory (VSTM) is compromised by information bottlenecks and visual interference between memorization and recall. Retro-cues, displayed after the offset of a memory stimulus and prior to the onset of a probe stimulus, indicate the test-item and improve performance in VSTM tasks. It has been proposed that retro-cues aid recall by transferring information from a high-capacity memory store into visual working memory (multiple-store hypothesis). Alternatively, retro-cues could aid recall by redistributing memory resources within the same (low-capacity) working memory store (single-store hypothesis). If retro-cues provide access to a memory store with a capacity exceeding the set size, then, given sufficient training in the use of the retro-cue, near-ceiling performance should be observed. To test this prediction, 10 observers each performed 12 hours across 8 sessions in a retro-cue change-detection task (40,000+ trials total). The results provided clear support for the single-store hypothesis: retro-cue benefits (difference between a condition with and without retro-cues) emerged after a few hundred trials and then remained constant throughout the testing sessions, consistently improving performance by two items, rather than reaching ceiling performance. Surprisingly, we also observed a general increase in performance throughout the experiment in conditions with and without retro-cues, calling into question the generalizability of change-detection tasks in assessing working memory capacity as a stable trait of an observer. Data and materials are available at osf.io/9xr82 and github.com/paulzerr/retrocues. In summary, the present findings suggest that retro-cues increase capacity estimates by redistributing memory resources across memoranda within a low-capacity working memory store.

Published as:

Zerr, P., Gayet, S., van den Esschert, F., Kappen, M., Olah, Z., & Van der Stigchel, S. (2021). The development of retro-cue benefits with extensive practice: Implications for capacity estimation and attentional states in visual working memory. *Memory & Cognition*, 1-14.

PZ, SG and SS designed the study concept. PZ programmed the experiment and wrote the manuscript. PZ, FE, MK, and ZO collected experiment data. PZ conducted the analyses. Critical revisions were provided by all co-authors

INTRODUCTION

Subjective perceptual experience suggests that the human visual system can represent many objects simultaneously. When visual input is no longer available, observers can maintain multiple objects in visual short-term memory (VSTM), which traditionally comprises visual working memory (VWM) and sensory memory (Baddeley & Hitch, 1974; Sperling, 1960; Averbach & Coriell, 1961; Table 1). A more recent conceptualization describes three states in VSTM: “activated” long-term memory (LTM; Table 1), a large capacity store, (2) a capacity-limited, attended “region of direct access” (VWM) and (3) a strongly attended, single item in the direct focus of attention (Oberauer, 2002; Oberauer & Hein, 2012).

Table 1 Names, abbreviations, and definitions of (hypothesized) short-term memory stores discussed in this article

Term	Definition
Visual short-term memory (VSTM)	An overarching term, describing any combination of memory systems that allow for maintaining visual information for a short amount of time.
Sensory memory	A short-lived (100 ms to 2 s), highly fragile, high-capacity buffer of visual information, which is unattended and not available for direct report but contains cueable (bound) objects.
Iconic memory	Sensory memory as described by Sperling (1960). Extremely short lifetime (up to 500 ms) and near-unlimited capacity.
Fragile memory	Sensory memory as described by Sligte, Scholte, and Lamme (2008). Moderate lifetime (up to 2 s) and moderate capacity (16 items).
Visual working memory (VWM)	A highly capacity-limited memory system, which relies on a limited resource (i.e., attention).
Long-term memory (LTM)	Unattended, near-unlimited capacity storage system. According to Oberauer (2002), some portion of it (activated LTM) can supplement working memory performance.

Note. The definitions provided in this glossary are our best attempt of paraphrasing the current status quo in the literature regarding these different memory stores. Whereas some of these memory stores have received considerable empirical support, it should be noted that the existence of others is not (yet) established.

While the structure and substrates of these memory systems are not yet fully understood, VWM is generally recognized as highly limited in capacity (e.g. “the magic number four”; Cowan, 2001; Luck & Vogel, 1997). Optimal operation of such a resource-constrained system requires a flexible allocation of a limited memory resource (i.e. attention) to prioritize task-relevant items at the expense of task-irrelevant items. Here we define “attention” as a cognitive resource that can be deployed across a very limited amount of information simultaneously. As a result, items in VWM are not always homogeneously represented. Attention and VWM are closely related and considered to represent a common neural resource (LaBar et al., 1999; Awh & Jonides, 2001; Cowan, 2001; Mayer et al., 2007). When studied at low set sizes, attended memory items differ from unattended (or less attended) memory items in several ways.

Attended, but not unattended, items appear to be maintained in sensory cortices in addition to parietal and frontal areas (Christophel et al., 2018; Sahan et al., 2019), interact with incoming sensory information (van Moorselaar et al., 2015; van Loon et al., 2017), and are maintained through persistent neural activity (i.e. sustained firing; Manohar et al., 2019; see also Stokes, 2015; Wolff et al., 2017; Wolff et al., 2015). Importantly, unattended items are more susceptible to perceptual interference than attended items, i.e. they are more likely to be erased by incoming sensory information (Makovski & Jiang, 2007; Matsukura et al., 2007; Makovski et al., 2010; Pinto et al., 2013; Souza et al., 2016). It has been proposed that attended and unattended items in VWM are represented in a qualitatively different state, such as neural spiking activity-based versus activity silent working memory storage (Stokes, 2015, Wolff et al., 2017; Wolff et al. 2015). Others have argued that the difference between attended and unattended items is merely quantitative, with attended and unattended items both reflecting activity-based storage (e.g., Schneegans & Bays, 2017; Rademakers & Serences, 2017). Common to all views, however, is that items within VWM are not represented homogenously, as different amounts of attentional resources can be allocated to different items.

It is widely accepted that observers can flexibly reallocate attentional resources within VWM, to prioritize task-relevant objects for recall and cognitive manipulation. Due to the inherent capacity-limitation of VWM, it would be beneficial if observers could also access latent items in separate large-capacity memory stores, from which information can be retrieved by retrospectively attending and transferring items to accessible VWM. Both LTM and sensory memory would represent candidates for such a qualitatively different, large-capacity memory store as these do not have the same resource limitation as VWM (i.e. attention). It remains unknown, however, whether retrospective allocation of attention allows observers to retrieve items from separate high-capacity memory stores into VWM.

The **retro-cue** paradigm (Landmann et al., 2003; Griffin & Nobre, 2003) emerged as a powerful experimental tool to study how an unattended item that was initially stored in a weak state can later be re-prioritized (i.e. attended; Lewis-Peacock et al., 2012; Sprague et al., 2016; LaRocque et al., 2017; van Loon et al., 2018; de Vries et al., 2018). A retro-cue is presented after offset of the to-be-memorized stimuli and predicts which item will be tested in the upcoming memory task. This cue allows observers to retrospectively prioritize an item in memory, increasing its likelihood of recall compared to a **post-cue**

condition, in which the test item is designated at the onset of the memory task display. Importantly, retro-cue paradigms, as well as partial-report paradigms, demonstrate that more information is stored in memory than is readily available for report in typical change-detection paradigms.

Sperling's partial-report experiments already established that more items can be reported if the subset of items that will be tested is made known during the retention interval. A distinction emerged between sensory memory as a high-capacity, unstable, short-lived, and implicit memory system (i.e. iconic memory and informational persistence; Coltheart, 1980; Pratte, 2018) and VWM as a strongly capacity-limited but stable and reportable memory system. Strikingly, the retro-cue paradigm and the partial-report paradigm are conceptually very similar as both improve memory report by presenting a cue prior to the memory probe. However, retro-cue effects are widely assumed to operate within VWM (for a review of contemporary explanations of the retro-cue effect see Souza & Oberauer, 2016). On the other hand, Sligte et al. (2008) proposed that retro-cues can access a high-capacity VSTM store that is distinct from VWM: Fragile memory. They described fragile memory as a form of sensory memory (Table 1), characterized by a high capacity, a lifetime of several seconds, high susceptibility to visual interference (Pinto et al, 2013) and not being limited by the amount of attention available for distribution among memory items (Vandenbroucke et al., 2011). In this view, the visual system maintains a high-capacity buffer of location-feature bound objects (Pinto et al., 2013), which can be brought into more stable VWM for cognitive manipulation and retrieval (Vandenbroucke et al., 2011; Vandenbroucke et al., 2015; Sligte et al., 2011). The retro-cue paradigm would then, in the same way as partial-report cues, constitute a potent way of experimentally inciting participants to transfer items that were initially stored in an unreportable state into VWM for recall and manipulation. Both sensory memory and LTM would fit the description of a large-capacity memory system that is not limited by the same memory resource as VWM. In this study, we test the hypothesis that retro-cues are able to access information in a separate high capacity store, such as activated LTM (i.e., Oberauer, 2002), or sensory memory (i.e., Sligte et al., 2008).

The view that retro-cues access a separate high-capacity memory store predicts that an observer's VSTM capacity is mostly limited by their proficiency in utilizing the retro-cue, and not by the lower capacity limits of VWM (Figure 1). Indeed, when presenting observers with set sizes far above typical VWM capacity in a retro-cue change-detection paradigm, Sligte et al. (2008) observed dramatically high-capacity estimates. In the condition with the largest set size, observers were, on average, able to report about 16 items from a memory array of 32 items, which is remarkably higher than the typical 3-4 item limit associated with VWM (e.g. Luck & Vogel, 1997). In contrast, Cowan's k (Cowan, 2001), a measure of working memory capacity, remained stable at around 4 items in the post-cue condition irrespective of set size. They discovered that k scaled with set size in the retro-cue condition. It should be noted, that k intends to provide a measure of working memory capacity which does not scale with set size (e.g. Rouder et al., 2011), which is the case in change-detection paradigms without retro-cues. Sligte et al. (2008) regarded the observed high capacity estimates for the existence of a high-capacity sensory memory store. It is indeed an intriguing hypothesis that the visual system would be able to retrieve information from a memory system (i.e. sensory memory or LTM) that is not subject to the resource-limitations of VWM (i.e. attention). Indeed, in line with the view that observers can access more information in visual memory than estimated in typical working memory tasks, other researchers found evidence that observers were able to retain and access 30 images in a rapid serial visual presentation paradigm (Ansgar & Potter, 2014).

The interpretation that the retro-cue effect reflects access to a high-capacity memory system has been questioned by some researchers (Matsukura & Hollingworth, 2011; Makovski, 2012; Robinson & Irwin, 2019). There are several competing (and partially complementary) hypotheses to explain the retro-cue effect (Souza & Oberauer, 2016), most prominently a shift of attention within VWM representations to protect items from visual interference. Matsukura and Hollingworth (2011) argued that the high capacities observed by Sligte et al. (2008) are explained by a long practice session preceding the experiment. They suggested that changes in processes unrelated to memory capacity, e.g. the efficiency of perceptual processing, memory encoding, maintenance, comparison processes or involvement of long-term memory could account for an increase in performance in memory tasks. In our experience (e.g., Zerr et al., 2017), it indeed typically takes many trials for a retro-cue benefit (the difference between retro-cue and post-cue condition performance) to emerge. Importantly,

however, while performance in any task is expected to increase with practice, participants in the study by Sligte et al. (2008) performed just as many trials in the post-cue condition as in the retro-cue condition, yet these high capacity estimates were only observed in the retro-cue condition. Consequently, if practice effects underlie these high capacities, they would be related to learning how to use the retro-cue and should differentially apply to the retro-cue condition and the post-cue condition, rather than reflect increased proficiency in e.g. the encoding of the memory array or the parsing of the probe array, which should influence performance in both conditions similarly. Moreover, the observation that extensive practice is required for a retro-cue benefit to emerge, does not exclude the possibility that retro-cues provide access to a high capacity memory store: while internal shifts of attention are commonplace in real life, a retro-cue that allows for manipulating these internal shifts of attention experimentally (Landmann et al., 2003; Griffin & Nobre, 2003) is an unnatural stimulus, which requires extensive practice to be optimally used. Specific aspects of the paradigm that are impacted by practice, however, depend on whether or not retro-cues allow observers to access a high capacity memory store that is distinct from VWM. Accordingly, differences between the learning curves in the post-cue and retro-cue conditions may be informative of potential qualitative differences between memory stores that are accessed by shifts of attention.

Studies using the retro-cue paradigm (including Sligte et al., 2008), typically report results as an average collapsed across time. It therefore remains unclear if learning to use the retro-cue is a slow or rapid process, whether the retro-cue benefit continues to increase indefinitely with more practice, and whether training increases performance in both conditions simultaneously. Observers in Sligte et al. (2008) reached ceiling performance in some set size conditions (4 and 8), but not others (16 and 32). It is unclear what caused this pattern of results and whether observers could have reached ceiling performance in all set sizes given enough practice. In the current study, we measured memory capacity estimates throughout an extensive training period, to distinguish between the patterns of training effects that would provide evidence for or against the hypothesis that retro-cues access a high-capacity memory store. We consider two scenarios describing how extensive training might affect memory performance in the post-cue and retro-cue conditions over time (see also Figure 1):

H₁: Retro-cues access a mechanistically and functionally distinct high-capacity memory store (i.e. sensory memory or LTM), which is not subject to the attention-based capacity limitations of VWM. With practice, observers get better at utilizing the retro-cue and bring items into robust VWM for retrieval. This scenario predicts that performance in the retro-cue condition increases continuously (relative to the post-cue condition) until capacity estimates are reached that far exceed traditional VWM capacity (i.e. multiples thereof). Traditional VWM capacity is reflected in performance in the post-cue condition. If the capacity of a memory system that can be accessed through a retro-cue exceeds the set size, then, given enough practice, all (or most) items should become reportable.

H₂: Retro-cues operate within a single memory store (VWM) and with practice, observers get better at using the retro-cue to redistribute attentional resources and prioritize less attended items in VWM and increase the probability of retrieval. This scenario predicts that after its initial emergence, the retro-cue benefit stabilizes once the limited capacity of VWM, including less prioritized items, is reached. This capacity estimate would be expected to be not much larger than what is observed in the post-cue condition and certainly smaller than the current set size of 12 items.

We investigated how memory performance in the post-cue and retro-cue conditions develops over time to discern between the two scenarios described above. To this end, ten observers performed a retro-cue change-detection task with a set size of twelve items for twelve hours over the course of eight sessions. This also ensured that observers reached the maximum retro-cue benefit they could achieve.

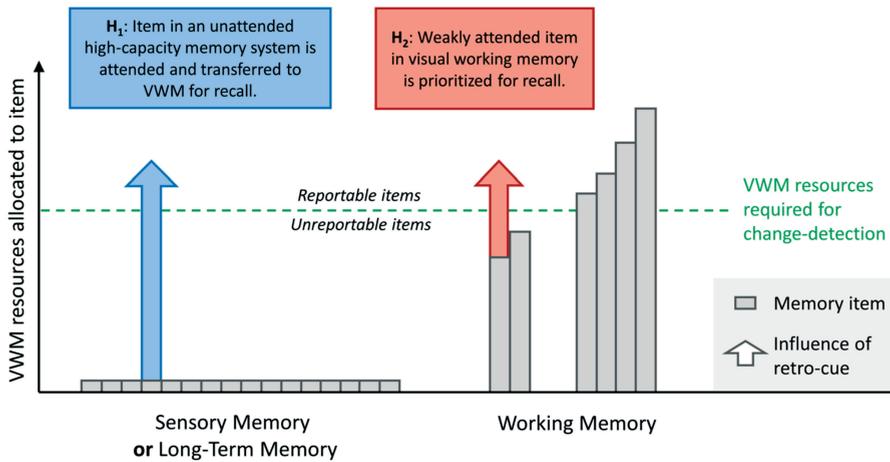


Figure 1: Graphical representation of the investigated hypotheses, which depicts two different ways in which information could be maintained in the visual system, and how reallocation of VWM resources following a retro-cue could act on these representations to improve performance in a change detection task. The y-axis represents the amount of attention (i.e. the memory resource underlying VWM) an item received. The bars on the left represent memory items in sensory memory or LTM, which can be cued (and accessed) with a retro-cue, but which are not subject to the same resource limitation as VWM. The bars on the right represent weakly and strongly attended items represented in low-capacity VWM. The green, dashed line represents the amount of VWM resources required for a memory object to be reportable in a change-detection task. A weakly attended item in VWM would not be reportable due to visual interference by the memory probe. The arrows represent how retro-cues would act on the different memory representations and increase capacity estimations under the multiple-store hypothesis (H₁; blue arrow) and single-store hypothesis (H₂; red arrow).

METHODS

Participants

Ten Utrecht University students (7 female; ages 19-36) took part in the experiment. This is the same number of participants that were tested in Experiment 1 in Sligte et al. (2008). All reported normal or corrected to normal vision and no psychiatric diagnosis.

Participants gave informed consent, and the study was approved by the local ethics committee. Participants received financial compensation or participant hours (course credits) for their time.

Setup and Stimuli

Stimuli were presented in a dark room on an ASUS PG278q LCD monitor with a display area of 60 x 34 cm with a resolution of 2560 x 1440 px at a refresh rate of 100 Hz and response time of 1 ms. Eye movements were monocularly recorded at 1000 Hz on an EyeLink1000 eye tracker (SR Research Ltd, Canada). Participants were seated on an adjustable chair with their head placed on a chin rest 65 cm in front of the screen. Stimuli were presented in Matlab 2015a and Psychtoolbox 3 (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007).

The trial layout is visualized in Figure 2. Each trial began with a blue fixation dot (0.5 dva; degrees visual angle), centered on the screen on a dark grey background and 12 light grey placeholder dots (0.23 dva), indicating the location of the upcoming memory items. Upon pressing the spacebar, the fixation dot turned red and after 500 ms, 12 red bars were presented as memory stimuli for 500 ms, randomly oriented in one of four possible orientations (0°, 45°, 90°, 135°). Exact stimulus positions can be extracted from the second upper panel of Figure 2.

In the **retro-cue** condition, after a 1000ms blank interval, during which only the fixation and placeholder dots were visible, a cue consisting of a white circle (0.6 dva) was presented for 250 ms at one of the 12 stimulus locations, followed by another 1000 ms blank interval, followed by the memory probe display. The retro-cue was always valid. In the **post-cue** condition, the memory display was presented immediately after the first 1000 ms blank interval.

The memory probe display consisted of 12 red, randomly oriented bars with a white circle (0.6 dva) indicating the test item, which either had the same orientation as in the memory stimulus (50% of trials) or was rotated by 90°. Participants pressed a key to indicate whether or not the test item had changed. The memory display remained visible until participants gave a response.

Procedure

Participants took part in 8 experimental sessions of 1.5 h each. The first five sessions took place Monday through Friday of the first week and the last three sessions Monday through Wednesday of the second week. At the beginning of the first session, participants were familiarized with the task during the course of 30 instruction trials (15 per condition), which were not included in the analysis. Instruction trials had a set size of 8 (instead of 12) and verbal instructions were provided by the experimenter. Every 8 trials contained 4 post-cue and 4 retro-cue trials presented in random order.

Participants were instructed to fixate on the dot in the center for the duration of a trial until they gave a response. If a participant's gaze deviated 2.5 dva from the central fixation dot, or if the participant blinked, the fixation dot turned into a blue "x", and the trial was aborted and repeated later. This was done to ensure that observers maintained central fixation during the crucial stages of the memory paradigm.

To ensure that the stimuli elicited no retinal afterimages the gray value of the screen background was calibrated to be perceptually isoluminant to the red stimuli for every participant at the start of each session. This was done by means of heterochromatic flicker photometry (Ives, 1912; Wagner, & Boynton, 1972). In addition, a full-screen colour noise mask was presented for 60ms (6 noise frames of 10ms) following the offset of the memory stimuli. Prior to the experiment, participants chose one of six cartoon characters which would "evolve" into another version of that character for each additional 5% accuracy the participants achieved. This was implemented to keep participants motivated and was well appreciated. After every 20 trials a progress graph was displayed that plotted correct and incorrect responses over time, with diagonal lines indicating participant's cumulative accuracy during the current experimental session. We have included a depiction of a progress screen in the online materials.

Participants received auditory feedback after each response, indicating a correct (low pitched sound; 300 Hz) or incorrect response (high pitched sound; 500 Hz). Auditory feedback was provided to facilitate task learning in the observers. Observers were encouraged to take short breaks every 10-15 minutes, after which the eye tracker was re-calibrated.

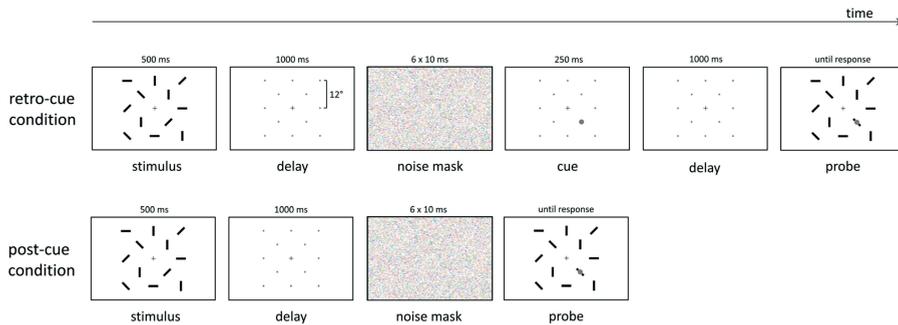


Figure 2: Trial layout. Sizes and colors are stylized in favor of visualization. First, a memory stimulus is displayed, consisting of 12 bars in one of four orientations. After a delay period, during which only the fixation dot and 12 placeholder dots are visible, a noise mask is presented. In the retro-cue condition a cue in the form of a small circle is then shown on one of the placeholder dots, which indicates the location of the upcoming target to be recalled. This is followed by a delay period and a memory probe display, consisting of the same 12 oriented bars, one of which has a 50% chance of being rotated by 90°. This target is indicated by a small circle, identical to the retro-cue. In the post-cue condition, instead of a cue display, the memory probe display is shown immediately after the noise mask, and simultaneously with the probe array.

Several differences between the present study and the paradigm used in Sligte et al. (2008) exist. Firstly, we use a noise mask rather than a light flash to mask retinal afterimages, because we found that this type of mask is less intrusive for observers than a bright flash of light after every trial, considering that they perform the task for 12 hours in total. The noise mask is as effective as a light flash, and, like the light flash, does not interfere with reportable sensory memory representations as previous research has established that representations that can be retrieved via the retro-cue are only susceptible to masks that resemble the memory stimulus (Pinto, 2013). Secondly, we use 12 items instead of varying set sizes. We test a specific prediction from Sligte et al. (2008): If retro-cues can access a memory system with a capacity of at least 16 items given a set size of 32, then, with enough practice, observers should reach ceiling performance with a set size of 12. This set size was chosen to be considerably above the “magic number four” (Cowan, 2001), while remaining well below the largest item-capacity observed in Sligte et al. (2008). Thirdly, while Sligte et al. (2008) used a block design for post-cue and retro-cue conditions, we interleaved trials in order to observe performance in both conditions evolve continuously over time (see

Setup and Stimuli). Importantly, these differences affect both conditions equally.

Data Analysis

In the figures and analyses, trial number indicates “trial time”, in which both conditions are assumed to proceed in parallel. Thus, trial time 100 is the 100th successive post-cue trial or the 100th successive retro-cue trial. Ten observers performed 42550 trials in total, 21275 per condition. Trials per observer and condition ranged between 1587 and 2502. Observers completed different amount of trials as they moved at different speeds. The analyses reported here were performed for trial time 1 to 1587, which is the last trial time that contains data from all ten observers.

To estimate performance (probability of a correct response) over time, a continuous accuracy score per trial (percent correct) was obtained by computing a moving average over binary performance trial data (correct/incorrect) from each observer. The centered moving average window was shrunk when reaching the start and end of the data array, such that the filtered signal at e.g. trial time 20 represents the average of 20 trials to the left and 20 trials to the right. This estimation, however, is dependent on the window size used in the moving average, which acts as a low-pass filter. The signal is therefore a function of window size: small window sizes reveal high frequency fluctuations in performance, whereas large window sizes reveal low frequency fluctuations in performance. Since the frequency of relevant changes in performance over time (learning curves) is undetermined, we present analyses for different window sizes in the online materials. Very small window sizes mostly pick up on noise (fast performance fluctuations are most likely related to variations in attention) and very large window sizes provide no additional benefit in noise reduction. The primary result of the present paper (i.e., the qualitative development of the retro-cue benefit over time) was not dependent on window size. However, the estimated time point at which the retro-cue benefit first reliably emerged, and the estimated time point at which it plateaued did vary as a function of window size. As optimal visualization duals optimal analysis, and for simplicity, we present average performance per condition in Figure 3 and report model comparisons for window size 200 only and provide results for all window sizes in the online materials.

The results of the moving window averaging approach can be interpreted by considering the filtered data as representing one participant's success rate at a given trial time in the experiment, when averaging over a number of trials equal to window size. For instance, in our data, observer 1 has an accuracy of 0.81 at trial 200 when using a moving average window size of 100. This means that if the participant performs 100 trials in an experiment after training for 150 trials, we expect a resulting percentage correct of 81% for those trials. Note that the continuous accuracy values from trial time one to the trial time that is equal to one half of the window size are subject to shrinking window sizes (see previous paragraph) and thus, datapoints at the very beginning of the array should be interpreted with caution. The same is true for the datapoints at the end of the array.

Using the filtered data from the ten observers, Bayesian estimation performed in JAGS (Plummer, 2003), using the *matjags* Matlab interface (Steyvers, 2011), produced group means and variance for each trial and condition. These group means (thick lines in Figure 3A) can be considered a hyperparameter of the population distribution which is generating these success rates. Cowan's k was approximated based on accuracy scores using equation (1) below

$$k = (2 \cdot \text{accuracy} - 1) \cdot \text{set size} \tag{1}$$

The retro-cue benefit was subsequently calculated by subtracting the estimated group means in the post-cue condition from those in the retro-cue condition. A linear, exponential, and logistic function (Table 2; equations 4-6) was fitted on the retro-cue benefit through least squares regression in Matlab. The Bayesian Information Criterion (BIC; Schwartz, 1978) of the fits were compared to determine the most likely model underlying the development of the retro-cue benefit over time. BIC was calculated using equation (2) below, where LL is the log-likelihood of the model, N is the number of datapoints, and k is the number of parameters in the model. Larger (negative) BIC's indicate a better model fit.

$$\text{BIC} = -2 * LL + \ln(N) * k \tag{2}$$

In addition, we estimated the standardized effect size between conditions for each trial using equation (3) below, and quantified group level evidence as Bayes factors in favor of an effect size greater than zero using a *Student's t* prior with $df = 1$ (i.e. a standard Cauchy prior; Lee & Wagenmakers, 2014; p. 124).

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}} \quad (3)$$

We welcome the re-use of the data, stimulus materials, and analysis scripts, which are available at osf.io/9xr82 and github.com/paulzerr/retrocues.

Table 2 Model equations

Equation	Model	$f(x_i)=$	Free parameters	
(4)	Linear	$m \cdot x_i + c$	m, c	slope, intercept
(5)	Exponential	$a \cdot e^{-k \cdot x_i} + c$	k, a, c	exponent, scaling factor, intercept
(6)	Logistic	$\frac{l_1 + (l_0 - l_1)}{1 + \left(\frac{x_i}{x_0}\right)^k}$	k, x_0, l_0, l_1	exponent, inflection point, lower asymptote, upper asymptote

RESULTS

Retro-cue benefit did not increase indefinitely with practice

Figure 3 A & B clearly demonstrate that the retro-cue benefit did not continue to increase over time after its initial emergence. A logistic function (BIC = - 9799) provided a better model fit than a linear (BIC = - 8228), or exponential function (BIC = - 9781). This result also becomes immediately apparent to the “naked eye” when observing Figure 3B. The logistic model suggests that participants were initially not better at retro-cue trials than at post-cue trials, then rapidly learned to use the retro-cue, and the retro-cue benefit subsequently reached a stable level. The logistic and exponential models converged on a maximum difference in capacity of $\Delta k = 2$

items, or a difference in accuracy of 8.3% (i.e. the asymptote of the functions).

When removing one subject at random from the analysis, BIC's for linear fits range from -7793 to -8597, exponential fits range from -9056 to -10056 and logistic fits range from -9075 to -10140. This means that in each permutation, the exponential or logistic model (both indicating that the retro-cue benefit reaches a plateau) is preferred over the linear model (indicating a continuous increase). The logistic model is preferred over the exponential model in 7 out of 10 permutations.

In the first derivative of the logistic fit (i.e. the function of change in the magnitude of the retro-cue benefit) the point when the retro-cue benefit changes less than 0.01% per trial (i.e. less than 1% per 100 trials) is reached at trial time 291, the point when the retro-cue benefit changes less than 0.001% per trial (i.e. less than 0.1% per 100 trials) is reached at trial time 535. This provides an indication for when the retro-cue benefit stops to increase in a meaningful way in the present data.

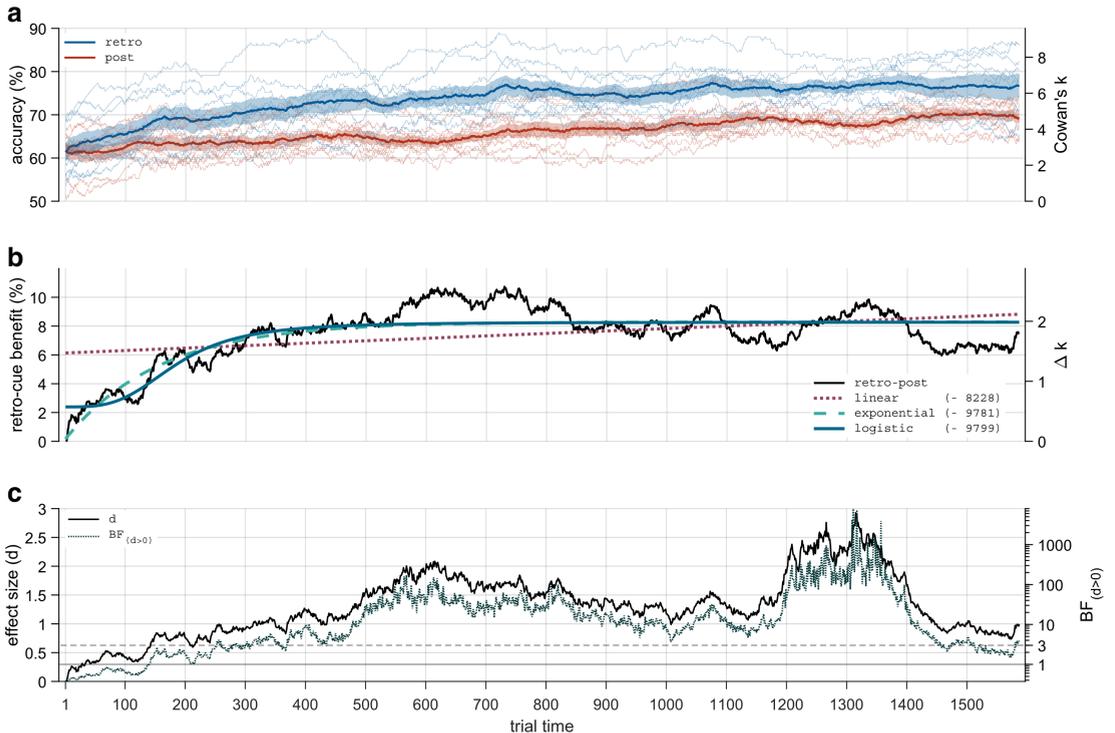


Figure 3. Main results for a moving average window size of 200 trials. Result graphs for a wide range of window sizes can be obtained in the online materials at osf.io/9xr82 and an animated version (gif) at osf.io/wqz8g. **A)** Accuracy in the post-cue and retro-cue conditions as moving averages. Thin lines indicate results from individual subjects. Thick lines indicate the estimated mean accuracy for each trial time. Shaded areas indicate estimated standard error. **B)** Retro-cue benefit (retro-cue accuracy minus post-cue accuracy) and model fits. **C)** Effect size and evidence for an effect size greater than zero for each trial time. The solid horizontal line indicates $BF = 1$ and the dashed horizontal line indicates $BF = 3$, a popular threshold for “evidence worth considering”.

Figure 4 displays data of individual observers as percent correct. Exploratory analyses revealed a mean change of the retro-cue benefit between the second half of one session and the first half of the following session of 0.002 (or 0.2% accuracy) across participants (SE: 0.01), indicating transfer of learning between sessions but no evidence for sleep consolidation effects on the group level.

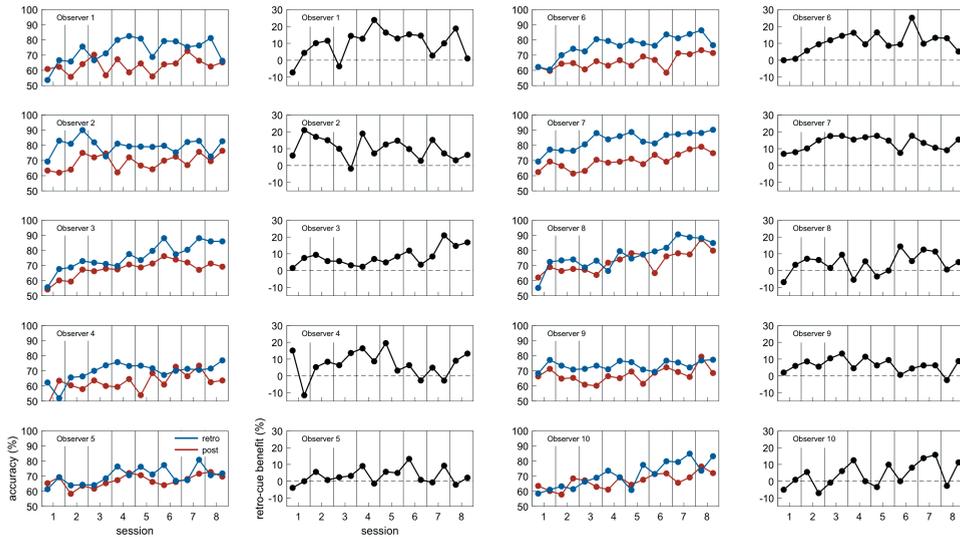


Figure 4: Individual observer data as percent correct, split into two bins per experimental session. All observers demonstrated a retro-cue benefit after some time.

Emergence of the retro-cue benefit

As described in the *Data Analysis* section, the result of the moving average transformation plotted in Figure 3A has a direct and meaningful interpretation: an accuracy value at trial N represents the proportion of correct trials an observer obtained across a number of trials that is retro equal to the window size used after $N - (\text{window size} / 2)$ trials of training. This translates to the level of evidence a researcher might obtain after a given amount of training and when using a given number of trials per condition. As can be seen in Figure 3C, evidence reliably crosses the popular $BF = 3$ threshold around trial time 300. Result patterns that are obtained when including different amounts of trials per condition can be observed in the online materials.

Accuracy continued to increase in both conditions

Notably, average accuracy continued to increase in both the post-cue and the retro-cue conditions (see Figure 3 A). An exploratory analysis indicated that linear models fitted from trial 500 to 1587 revealed slopes greater than zero ($m_{\text{post}} = 6.2 \cdot 10^{-5}$, 95% CI = $[6.0 \cdot 10^{-5}, 6.3 \cdot 10^{-5}]$, or 0.62% increase per 100 trials; $m_{\text{retro}} = 3.7 \cdot 10^{-5}$, 95% CI = $[3.5 \cdot 10^{-5}, 3.8 \cdot 10^{-5}]$, or 0.37% increase per 100 trials). This means that, at 1587 trials, observers' average performance was still increasing in both the post-cue and the retro-cue conditions. A linear model fitted from trial 500 to 1587 for the difference between conditions (i.e. the retro-cue benefit) revealed a slightly negative slope ($m_{\text{retro-post}} = -2.5 \cdot 10^{-5}$, 95% CI = $[-2.7 \cdot 10^{-5}, -2.3 \cdot 10^{-5}]$, or a 0.25% decrease per 100 trials). This provides further evidence that retro-cue benefits did not continue to increase indefinitely, and if anything, slightly decreased over time with excessive training.

DISCUSSION

Retro-cues allow observers to reallocate attention within the contents of VSTM to prioritize an item prior to a memory task and increase its probability of recall. While previous studies reported performance averaged across time, we investigated the learning curves of ten human observers performing a retro-cue change-detection task for 12 hours each. We compared the development of accuracy over time between retro-cue and post-cue conditions to discern qualitative and quantitative differences in learning rates, which could reveal differences in VSTM mechanisms involved in the two conditions.

No evidence that retro-cues access a separate high-capacity store

If retro-cues enable access to a high-capacity VSTM store that is able to hold 16 items or more (Sligte et al., 2008), then with enough practice, observers should be able to reach near ceiling performance, given our current set size of 12 items. However, in our experiment, the difference in performance between the retro-cue and post-cue

conditions quickly stabilized at about $\Delta k = 2$ items (or a difference of 8.3% accuracy). Specifically, this retro-cue benefit plateaued after a few hundred trials. At that point, performance had reached a capacity of $k = 4$ in the post-cue condition and a capacity of $k = 6$ in the retro-cue condition. This result does not support the hypothesis that retro-cues allow observers to access high-capacity sensory memory or LTM (H_1), as it would predict that in a memory task with a set size of 12 items, the retro-cue benefit continues to increase until observers reach an estimated working memory capacity of close to $k = 12$ items because sensory memory and LTM should not be subject to the same resource bottleneck (i.e. attention) as VWM. Instead, the data show that the retro-cue consistently adds 2 items to stable working memory capacity. This result favors the hypothesis that retro-cues allow observers to access a portion of VSTM which received sufficient attention to be encoded into capacity-limited VWM, but not enough attention to be reportable in a change-detection task without retro-cues (H_2). This implies that VWM contains more information than estimated in a task without retro-cues, or in other words: that some items in this memory store are encoded less robustly and can be prioritized retrospectively.

On the other hand, it has been observed in neuroimaging results that unattended items that could not be decoded during the retention interval, were once again decodable after an attentional shift induced by the retro-cue (e.g., Christophel et al., 2018; Sahan et al., 2019) and it has been suggested that objects in VSTM can exist in activity-silent states (e.g. Stokes, 2015, Wolff et al., 2017; Wolff et al. 2015). Such a qualitatively different neural storage implementation would support the idea of a categorically different memory state for attended and unattended items, as also proposed by the multiple-store hypothesis. In response to this interpretation, however, it has been suggested that unattended memory items are not stored in a categorically different way, but that the neural activity elicited by unattended items persists, albeit with an amplitude too small to be picked up by the imaging technique (e.g. Schneegans & Bays, 2017; Rademakers & Serences, 2017). The present results also speak more strongly for quantitatively, rather than qualitatively, different states within VWM.

Behaviorally, Makovski (2012) also concluded support for the single-store hypothesis from the finding that a retro-cue benefit exists even when the retro-cue is displayed after visual interference. However, the authors used a set size of only four items, which may be

insufficient to investigate a high-capacity system. Further evidence for a single store comes from Robinson and Irwin (2019), who used a state-trace analysis to assess the dimensionality of VSTM and concluded that the results were more parsimonious with the single-store hypothesis.

The activated long-term memory (LTM) hypothesis of working memory (Cantor & Engle, 1993; Oberauer, 2002; Ruchkin et al., 2003; Öztekin et al., 2010) describes a three-state model of working memory: (1) activated LTM, which can keep task-relevant information such as feature-spaces available and is not capacity- or attention-limited, (2) a “region of direct access”, which is closely related to traditional VWM, capacity-limited and represents a “broad focus” of attention, and (3) a single item in the focus of attention. Within this framework, the present results are more parsimonious with the view that retro-cues prioritize (focus attention on) one item from the broadly attended “region of direct access” and protecting it from interference (i.e., H_2), rather than moving information from LTM into the “region of direct access” (i.e., H_1).

In conclusion, while there remains the possibility that some objects in VSTM are represented by a qualitatively different mechanism and that the visual system contains high-capacity information stores, the present findings suggest that retro-cues operate within VWM by redistributing attentional resources and prioritizing a relevant object (Nobre et al., 2003; Souza and Oberauer, 2016). This implies that retro-cue paradigms (or for that matter partial-report paradigms) may not be suitable to investigate the existence or functionality of high-capacity memory stores. Other methodologies, such as rapid serial visual presentation paradigms (Ansgar & Potter, 2014), may yet be able to reveal and investigate high-capacity VSTM. Our results further suggest that VWM has a larger capacity than estimated in change-detection tasks without retro-cues.

Continuous performance increase in both conditions

While the retro-cue benefit remained stable, accuracy continued to increase in both conditions throughout the experiment and this increase was still present after 12 hours spent on the task. While this parallel increase in both conditions provides further support for the single-store hypothesis, it is also surprising. Working memory capacity is often considered to be a stable trait of an observer (Xu et al., 2018), and is correlated with several personality traits, such

as fluid intelligence (e.g. Feldman et al., 2004). Since different observers respond differently to the same amount of practice, this is especially problematic when the estimated capacity in a change-detection task is taken as an absolute measure (e.g. trait of an individual) such as in correlational (e.g. Shipstead et al., 2012) and developmental studies (e.g. Riggs et al., 2006; Simmering, 2012), rather than a relative measure (e.g. within-subject comparisons). In addition, most studies investigating working memory capacity report performance collapsed across time. Meta-analyses cannot adequately compare such studies without taking into account the amount of practice observers received.

Furthermore, it is highly unlikely that an increase in performance over time represents an increase in actual memory capacity, at least not within 12 hours. One possibility is that observers learned to distribute their attention more evenly across the memory items during encoding. Observers may also have been relieved of crowding effects over time (Yashar et al., 2015). Matsukura and Hollingworth (2011) also pointed to changes in processes unrelated to memory capacity, e.g. the efficiency of perceptual processing, memory encoding, maintenance, comparison processes or involvement of long-term memory as potential explanations for a continuous increase in performance in change-detection tasks. They tested the latter hypothesis in a control experiment by comparing performance between using bars with either two or four possible orientations, under the assumption that relations between items (i.e. chunking) would be more easily stored when using two possible orientations (horizontal and vertical). They found no difference between these conditions, suggesting that context effects do not explain high capacity estimates. Sligte et al. (2008) also tested this possibility and found no difference in performance when a single item or all items were shown in the memory task display. However, other researchers did find evidence that ensemble representations may facilitate memory processes and indeed may inflate capacity estimates, albeit for complex objects (Brady & Alvarez, 2015). As such, it remains unclear what factors drive the continuous increase in performance over time that we observed in the current experiment.

The finding that performance continuously increased in the post-cue condition casts doubt on the usefulness of change-detection tasks in estimating VWM capacity of an observer as it may not be possible to generalize across experiments that are employing different amounts of practice as well as difficulties when investigating

individual differences. More generally, an experimental method that yields an ever-increasing measure of VWM capacity may not be well suited for measuring actual VWM capacity, which is typically considered stable over time.

Practice does not account for previously observed, very high Cowan's k estimates

It is important to point out that the continuous increase in performance, described above, was observed in the post-cue condition as well as the retro-cue condition. Matsukura and Hollingworth (2011) also reported a continuous increase in performance when two observers practiced a retro-cue change-detection task over the course of 80 minutes. From this, they concluded that extensive practice could account for the large performance differences between post-cue and retro-cue conditions observed by Sligte et al. (2008). However, Matsukura and Hollingworth (2011) did not include a post-cue control condition to ascertain that the increase in performance was specific to the retro-cue condition. The present data show that these practice effects are not specific to the retro-cue condition, and therefore cannot account for the retro-cue benefit and high VWM item capacities. More specifically, Figure 3B of the present paper clearly shows that increases in performance that are specific to the retro-cue effect conclude after a few hundred trials, as retro-cue benefits no longer increase (while performance in both conditions continues to increase in parallel). In other words, continued practice in the task increases performance in both conditions and does not selectively increase the retro-cue benefit. Matsukura and Hollingworth (2011) further critiqued the multiple-store hypothesis by suggesting that capacity estimates drop from 16 to 5-7 items when two instead of four possible orientations are used. However, these lower capacity estimates were observed in experiments with a set size of only 8 items. Thus, the extremely high item-capacity estimates observed by Sligte et al. (2008) remain intriguing and will require further investigation to be reconciled with the VSTM literature, including our study.

The limits of the magic number four and implications for item-limits in VWM

The item-based capacity measure Cowan's k (Cowan 2001) is intended to provide an estimate of the number of items that an observer can maintain in memory irrespective of set size. Indeed, k is

stable across set sizes (Rouder et al., 2011) and even across experiments (Xu et al., 2018) when no retro-cues are employed. However, with retro-cues, k (and the retro-cue benefit when expressed in k) does scale with set size, such as in the study by Sligte et al. (2008). This was also reported by Souza et al. (2014), albeit with a smaller set size of one to eight items. This means that when not using retro-cues, k underestimates the memory capacity of an observer, because more fragile, less attended memory items (which were nonetheless encoded into VWM), are not captured by the memory task. On the other hand, when retro-cues are employed, k is dependent on set size. While k seems to offer an intuitive interpretation of capacity as an item limit, these considerations call for caution in interpreting this metric as a proxy of VWM capacity.

These limitations likely stem from the assumption inherent in k that the underlying resource in VWM is composed of discrete slots. This view finds support in the following observation: When the results by Sligte et al. (2008; taken from Figure 2, page 3) are back-transformed from Cowan's k to percentage correct, performance for set sizes 4, 8, 16 and 32 dropped from about 88% to 74%, 62% and 60% in the post-cue condition, while performance in the retro-cue condition dropped from about 100% to 94%, 84% and 75%. Thus, the observation that the retro-cue benefit (and therefore VWM capacity) scales with set size no longer holds when the differences between conditions are expressed as percentage correct: 12%, 20%, 22% and 15%. Percentage correct can be interpreted as the average probability of any item to be recalled. A similar retro-cue benefit to recall probability across different set sizes suggests that retro-cues enable access to a fixed amount of a flexible resource (attention or the information carrying capacity of the VWM system; e.g. Bays & Husain, 2008; Schneegans & Bays, 2016) rather than a fixed number of item slots (e.g. Luck & Vogel, 2008). This view is also more supported by our present findings, which suggest that retro-cues re-distribute attention within VWM. For a discussion of retro-cue benefits in relation to discrete slot (e.g. Cowan's k) and flexible-resource models of capacity-limits in VWM that is beyond the scope of the present paper see Souza et al. (2014).

Large retro-cue benefits necessitate long practice sessions

An important conclusion from the present data is that observing large and reliable retro-cue benefits requires generous

practice before experimental data is collected. Large individual differences exist in the speed and extent of learning the usage of the retro-cue and a long training session before recording experimental data is highly recommended to avoid false nulls and especially if it is not the retro-cue benefit itself that is of interest but a modulation thereof. Furthermore, it should be avoided to collect experimental data on an effect that still continues to increase during the experiment session, especially if the rate of increase might differ between observers. Based on the current results, a good rule of thumb would be to include about 500 trials of practice before collecting experimental data for maximal effects.

Summary

Our results suggest that retro-cues operate within the contents of VWM and do not provide evidence that retro-cues access a high-capacity memory store that is independent of the resource-limitation of VWM (such as sensory memory or LTM). Additionally, the present data cast doubt on the usefulness of change-detection tasks and Cowan's k in estimating VWM capacity. Finally, we suggest that studies using change-detection retro-cue experiments employ generous amounts of practice trials before collecting experimental data.

REFERENCES

- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, 40, 309-328.
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in cognitive sciences*, 5(3), 119-126.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.
- Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 921.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10(4), 433-436.
- Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*

- Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, 21(4), 494-496.
- Coltheart, M. (1980). The persistences of vision. *Philosophical Transactions of the Royal Society of London: Series B*, 290, 269-294.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24, 87-185.
- de Vries, I. E., van Driel, J., Karacaoglu, M., & Olivers, C. N. (2018). Priority switches in visual working memory are supported by frontal delta and posterior alpha interactions. *Cerebral Cortex*, 28(11), 4090-4104.
- Feldman Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual difference in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, 130(4), 553-573. doi:10.1037/0033-2909.130.4.553.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15, 1176-1194. doi:10.1162/089892903322598139
- Ives, H.E. (1912). On heterochromatic photometry. *Philosophical Magazine U*, 845-853
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?.
- LaBar, K. S., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *Neuroimage*, 10(6), 695-704.
- Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149-164. doi:10.1016/S0042-6989(02)00402-9
- LaRocque, J. J., Riggall, A. C., Emrich, S. M., & Postle, B. R. (2017). Within-category decoding of information in different attentional states in short-term memory. *Cerebral Cortex*, 27(10), 4881-4890.
- Lee, M. D., & Wagenmakers, E. J. (2014). Bayesian cognitive modeling: A practical course. *Cambridge university press*.
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, 24(1), 61-79.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.
- Manohar, S. G., Zokaei, N., Fallon, S. J., Vogels, T., & Husain, M. (2019). Neural mechanisms of attending to items in working memory. *Neuroscience & Biobehavioral Reviews*.
- Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics*, 69, 1422-1434. doi:10.3758/BF03192957
- Matsukura, M., & Hollingworth, A. (2011). Does visual short-term memory have a high-capacity stage? *Psychonomic Bulletin & Review*, 18(6), 1098-1104. <https://doi.org/10.3758/s13423-011-0153-2>

- Makovski, T., Watson, L. M., Koutstaal, W., & Jiang, Y. V. (2010). Method matters: Systematic effects of testing procedure on visual working memory sensitivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1466–1479. doi:10.1037/a0020851
- Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect?. *Psychonomic bulletin & review*, 19(3), 470-476.
- Mayer, J. S., Bittner, R. A., Nikolić, D., Bledowski, C., Goebel, R., & Linden, D. E. (2007). Common neural substrates for visual working memory and attention. *Neuroimage*, 36(2), 441-453.
- Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 411.
- Öztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. *Psychological Science*, 21, 1123–1133.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10(4), 437-442.
- Pratte MS. Iconic memories die a sudden death. *Psychol Sci*. 2018;29(6):877-887. doi:10.1177/0956797617747118
- Pinto, Y., Sligte, I. G., Shapiro, K. L., & Lamme, V. A. (2013). Fragile visual short-term memory is an object-based and location-specific store. *Psychonomic bulletin & review*, 20(4), 732-739.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10, pp. 1-10).
- Rademaker, R. L., & Serences, J. T. (2017). Pinging the brain to reveal hidden memories. *Nature neuroscience*, 20(6), 767.
- Riggs, K. J., McTaggart, J., Simpson, A., & Freeman, R. P. (2006). Changes in the capacity of visual working memory in 5-to 10-year-olds. *Journal of experimental child psychology*, 95(1), 18-26.
- Robinson, M. M., & Irwin, D. E. (2019). Are there two visual short-term memory stores? A state-trace analysis. *Journal of Mathematical Psychology*, 90, 23–30. <https://doi.org/10.1016/j.jmp.2019.02.001>
- Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic bulletin & review*, 18(2), 324-330.
- Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, 26(6), 709-728.
- Sahan, M. I., Sheldon, A. D., & Postle, B. R. (2019). The Neural Consequences of Attentional Prioritization of Internal Representations in Visual Working Memory. *Journal of Cognitive Neuroscience*, 1-28.
- Schneegans, S., & Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *cortex*, 83, 181-193.

- Schneegans, S., & Bays, P. M. (2017). Restoration of fMRI decodability does not imply latent working memory states. *Journal of Cognitive Neuroscience*, 29(12), 1977-1994.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.
- Shipstead, Z., Redick, T. S., Hicks, K. L., & Engle, R. W. (2012). The scope and control of attention as separate aspects of working memory. *Memory*, 20(6), 608-628.
- Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of experimental child psychology*, 111(4), 695-707.
- Steyvers, M. (2011). MATJAGS 1.3: A Matlab interface for JAGS.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores?. *PLOS one*, 3(2).
- Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Scholte, H. S., & Lamme, V. A. (2011). Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia*, 49(6), 1578-1588.
- Souza, A.S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance*. doi:10.1037/a0036331
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74(Whole No. 498), 1-29.
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, 91(3), 694-707.
- Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702-712.
- Stokes, M. G. (2015). "Activity-silent" working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19, 394-405.
- van Loon, A. M., Olmos-Solis, K., & Olivers, C. N. (2017). Subtle eye movement metrics reveal task-relevant representations prior to visual search. *Journal of Vision*, 17(6), 13.
- van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *ELife*, 7, e38677.
- van Moorselaar, D., Olivers, C. N., Theeuwes, J., Lamme, V. A., & Sligte, I. G. (2015). Forgotten but not gone: Retro-cue costs and benefits in a double-cueing paradigm suggest multiple states in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1755.
- Vandenbroucke, A. R., Sligte, I. G., & Lamme, V. A. (2011). Manipulations of attention dissociate fragile visual short-term memory from visual working memory. *Neuropsychologia*, 49(6), 1559-1568.
- Vandenbroucke, A. R., Sligte, I. G., de Vries, J. G., Cohen, M. X., & Lamme, V. A. (2015). Neural correlates of visual short-term memory dissociate between fragile and working memory representations. *Journal of cognitive neuroscience*, 27(12), 2477-2490.

RESEARCH ARTICLES

- Wagner, G., & Boynton, R. M. (1972). Comparison of four methods of heterochromatic photometry. *JOSA*, 62(12), 1508-1515.
- Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in systems neuroscience*, 9, 123.
- Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature neuroscience*, 20(6), 864-871.
- Yashar, A., Chen, J., & Carrasco, M. (2015). Rapid and long-lasting reduction of crowding through training. *Journal of Vision*, 15(10), 15-15.
- Xu, Z., Adam, K. C. S., Fang, X., & Vogel, E. K. (2018). The reliability and stability of visual working memory capacity. *Behavior Research Methods*, 50(2), 576-588.
- Zerr, P., Gayet, S., Mulder, K., Pinto, Y., Sligte, I., & Van der Stigchel, S. (2017). Remapping high-capacity, pre-attentive, fragile sensory memory. *Scientific reports*, 7(1), 1-10.

2.4 Successful visually guided eye movements following sight restoration after congenital cataracts

ABSTRACT

Sensitive periods have previously been identified for several human visual system functions. Yet it is unknown to which degree the development of visually guided oculomotor control depends on early visual experience, for instance, whether and to which degree humans whose sight was restored after a transient period of congenital visual deprivation are able to conduct visually guided eye movements. In the present study we developed new calibration and analysis techniques for eye tracking data contaminated with pervasive nystagmus as is typical for this population. We investigated visually guided eye movements in sight recovery individuals with long periods of visual pattern deprivation (3-36 years) following birth due to congenital, dense, total, bilateral cataracts. As controls we assessed 1) individuals with nystagmus due to causes other than cataracts, 2) individuals with developmental cataracts after cataract removal, and 3) individuals with normal vision. Congenital cataract reversal individuals were able to perform visually guided gaze shifts, even when their blindness had lasted for decades. The typical extensive nystagmus of this group distorted eye movement trajectories, but measures of latency and accuracy were as expected from their prevailing nystagmus, that is, not worse than in the nystagmus control group. The present, first, quantitative study on the characteristics of oculomotor control in congenital cataract reversal individuals indicates a remarkable effectiveness of visually guided eye movements despite long lasting periods of visual deprivation.

Published as:

Zerr, P., Ossandón, J. P., Shareef, I., Van der Stigchel, S., Kekunnaya, R., & Röder, B. (2020). Successful visually guided eye movements following sight restoration after congenital cataracts. *Journal of Vision*, 20(7), 3-3

PZ, JO and BR designed the study concept. PZ programmed the experiment and wrote the manuscript. PZ and IS collected experiment data. PZ and JO conducted the analyses. Critical revisions were provided by all co-authors.

INTRODUCTION

Gaze orienting is crucial for visual perception. Eye movements allow us to foveate relevant events in the world and require a precise and efficient oculomotor system. Humans are not born with a fully functional oculomotor system, which is characterized by a protracted development (Luna et al., 2008). While visually guided saccades have been reported to occur as early as within the first month of life (Aslin & Salpatek, 1975), infants and children had less stable fixations, longer latencies and shorter saccade amplitudes than adults (Alahyane et al., 2016; Aslin & Salpatek, 1975; Cohen & Ross, 1978; Ross et al., 1994; Harris et al., 1993; Munoz et al., 1998; Irving et al., 2011). Eye movement studies in blind humans have demonstrated that while many late-blind individuals were able to voluntarily orient their eyes even after several years of blindness, congenitally blind adults with intact eyes were not (Leigh & Zee, 1980; Kämpf & Pieper, 1987; Hall et al., 2000; Schneider et al., 2013). These results suggest that the human oculomotor system does not properly develop in the total absence of visual input. However, it is yet unclear whether and to what degree eye movement control is possible in humans whose sight can be restored, that is, whether there is a sensitive period for the development of visually guided eye movements.

Individuals born with congenital, dense, total, bilateral cataracts provide a unique opportunity for research into the experience-dependent development of the human visual system as several of them do not receive patterned visual input until the cataracts are removed. While some visual system functions appear to develop normally after sight is restored, such as colour perception (Brenner et al., 1990; Pitchaimuthu et al., 2019; McKyton et al., 2015), retinotopy (Sourav et al., 2018) and biological motion detection (Bottari et al., 2015; Hadad et al., 2012), others do not, leading to lasting impairments such as reduced visual acuity (Elleberg et al., 1999; Kalia et al., 2014), impaired face processing (Le Grand et al., 2001; Putzar et al., 2010), reduced global motion perception (Bottari et al., 2018; Hadad et al., 2012) and compromised multisensory speech perception (Putzar et al., 2007, 2010). It could be hypothesized that these impairments stem from an impaired oculomotor control, preventing these individuals from systematically scanning the visual environment. However, the ability to perform visually guided eye movements after congenital visual deprivation has not yet been quantitatively investigated, that is, a systematic assessment of visually guided oculomotor control in cataract reversal individuals with visual

deprivation periods extending beyond the typical range of sensitive periods for development of visual system functions is missing. Studies on the effects of transient binocular visual deprivation in macaque monkeys (Carlson, 1990; Harwerth et al. 1991; Hyvärinen et al, 1981; Regal et al., 1976; Tusa et al. 2001) reported that after sight restoration the animals were able to visually orient their head and eyes to follow light and objects, but did not analyse oculomotor control in detail. The main reason for a lack of detailed data on eye movement control in cataract reversal individuals most likely is the extensive and permanent nystagmus these individuals develop if surgery was not performed within the first weeks of life. These involuntary oscillations of the eyes render the estimation of gaze position extremely difficult.

The present study developed a new approach to assess eye movements in cataract reversal individuals despite their prevailing extensive nystagmus and investigated to which extent human oculomotor control develops after a long, period of congenital blindness. Eye tracking was performed in a reactive saccade task and eye movement characteristics including 1) latency, 2) duration, 3) endpoint error, 4) saccade peak velocity and 5) saccade amplitude were assessed. The data of congenital cataract reversal individuals were compared to three control groups comprising developmental cataract reversal individuals, to assess effects of age at visual deprivation onset, individuals with nystagmus due to other reasons than cataracts, to estimate the effects of nystagmus on visually guided eye movements, and normally sighted controls.

METHODS

Participants

All participants were recruited and tested between November 2018 and February 2019 at the LV Prasad Eye Institute in Hyderabad, India. None of the participants had impairments in sensory systems other than vision, nor any known neurological disorder. One additional nystagmus control individual (see below) was excluded for excessive eye tracking signal artefacts due to partially closed eyes.

The cataract reversal group (CC) consisted of 9 subjects (3 female, mean age: 22.8 years, range: 8 - 44) with a history of congenital, dense, total, bilateral cataracts which were surgically

removed earliest at the age of three years (mean: 15.8 years, range: 3 - 36 years). The history of congenital, dense, total, bilateral cataracts was confirmed by medical records. In addition to the clinical diagnosis, factors such as vision loss, the density of the lenticular opacity, strabismus, presence of nystagmus, absence of fundus view prior to surgery, and a positive family history aided in the classification of CC participants. Pre-surgical visual acuity assessments were available in 6 of the 9 cases. Three CC individuals with a pre-surgery visual acuity between 1.4 to 1.48 logMAR units (see Table 1) had partially absorbed cataracts, which are regularly observed in congenital cataract patients. A visual acuity of 1.4 logMar is considered as severe blindness according to the ICD-10 (World Health Organization, 2004). CC individuals were tested at least one year after cataract removal surgery (mean: 7.0 years, range: 1 - 19). Mean post-surgical visual acuity was 0.94 logMAR (range: 0.40 - 1.30). All CC individuals suffered extensive nystagmus. Table 1 provides a detailed description of the CC individuals.

The nystagmus control group (NC) group consisted of 13 individuals with infantile nystagmus syndrome due to other reasons than cataracts and without a period of blindness. These individuals enabled a dissociation of the influence of nystagmus and early visual experience (3 female, mean age: 12.5 years, range: 7 - 28; mean visual acuity: 0.50 logMAR, range: 0.10 - 0.90).

The developmental cataract group (DC) consisted of 16 individuals (6 female, mean age: 15.3 years, range: 9 - 32). All DC individuals had a history of a transient period of bilateral cataracts later in life and had undergone the same surgical procedure as the CC individuals. They were tested at least one year after cataract removal surgery (mean: 9.0 years, range: 1 - 19). Mean post-surgical visual acuity was 0.20 logMAR (range: 0.00 - 1.00).

The sighted control group (SC) consisted of 13 individuals (3 female, mean age: 14.6 years, range: 10 - 18) with normal or corrected-to-normal vision. This group allowed us to establish typical eye movement parameters for healthy individuals.

Expenses associated with taking part in the study were reimbursed. Minors received a small present. The participants and/or their legal guardians provided written informed consent for taking part in the study. Participants and/or their legal guardians were informed about the study and received the instructions in one of the languages they were able to understand. The study was approved

by the local ethics board of the Faculty of Psychology and Human Movement Science (University of Hamburg, Germany) and the institutional ethics board of the LV Prasad Eye Institute (Hyderabad, India).

Setup

Eye movements were recorded binocularly at 500 Hz using an Eyelink1000 Plus eye tracker (SR Research Ltd, Canada). Subjects were seated in a darkened room and head constrained so that their eyes were at a distance of 60 cm from the screen. Stimuli were generated in MATLAB (The MathWorks, Natick, MA) using Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007) on a Windows 7 PC and presented on an Eizo FG2421 LCD monitor at a resolution of 1920 x 1080 @ 120 Hz.

Experimental Design

After giving instructions and performing eye tracker calibration, a grey dot (diameter = 1°; degrees visual angle) was shown on a black screen. After 2 - 4 s (uniform random sample) the dot jumped to a new location at a distance of 5 - 15° (uniform random sample) in a random direction from the previous dot location. Targets could appear in a subsection of the screen spanning 39 x 22°. Subjects were instructed to always fixate the dot and to follow the target when it changed position. Thus, the saccade target of a trial always became the fixation target of the next trial. To prevent fatigue and to minimize blinking, after every 10 trials (dot position changes) the dot disappeared. Subjects were instructed to close their eyes and after a few seconds to open them again. After the experimenter pressed a key the dot reappeared at the same location from which it had disappeared, and after 2 seconds jumped to the next location. After 99 dot jumps the experiment terminated.

Participant	Age, y	Gender	Age at surgery, y	Visual experience, mo	Presurgical visual acuity	Postsurgical visual acuity
CC-a	36	Male	23	161	Unknown	1.30
CC-b	17	Male	16	12	1.40 [*]	0.80
CC-c	16	Female	15	12	1.48 [*]	1.10
CC-d	13	Female	3	120	Unsteady fixation at light	0.40
CC-e	44	Male	36	102	Unknown	1.20
CC-f	8	Female	4	48	CF 0.5 m	0.70
CC-g	22	Male	17	58	CF 0.5 m	0.90
CC-h	32	Male	13	226	Unknown	1.00
CC-i	16	Male	15	15	1.48 [*]	1.10

Table 1. Characteristics of the CC individuals. *Notes:* Visual acuity is expressed in logMAR units and refers to the better eye. CF represents those able to count fingers; equivalence with logMAR acuity has been reported to be 1.7 to 2.0 at 30 cm (Schulze-Bonsel, Feltgen, Burau, Hansen, & Bach, 2006). ^{*}Subjects with documented absorbed cataracts in the latter part of their childhood.

Online eye-tracker calibration

The extensive nystagmus present in many of the subjects prohibited the use of default calibration procedures. Each subject was initially calibrated manually by selecting gaze samples from raw data recordings while a target was presented at one of following five locations: screen centre, 15° right and left of the centre and 8.5° above and below the centre. The mapping was determined by fitting the medians of the selected samples to the corresponding calibration targets using a second order polynomial function (Stampe, 1993).

Offline calibration and endpoint error

Optimal calibration coefficients were determined offline by minimising median endpoint error. Beginning with the initial online manual calibration, eye movement endpoints were computed by the procedure described in the section *Detection of visually guided eye movements* below. For each subject, the calibration coefficients that resulted in the lowest mean Euclidian endpoint error across all trials were determined by an implementation of the Nelder-Mead simplex optimization algorithm. These calibration coefficients were then used to recalibrate the raw gaze position data, from which new endpoints were computed. This procedure was repeated 30 times, from which the set of coefficients with the lowest resulting endpoint error were determined and used to calibrate the raw gaze position data for the analysis. Due to the random distribution of saccade targets across the screen, this method generated an optimal mapping function to minimize calibration error. This offline calibration method would however be insensitive to directionally systematic endpoint errors, such as for example always looking to the left of a target dot. Such

constant errors were not of interest in the present study and do not affect the interpretation of the extracted eye movement parameters.

The calibration procedure was 11-fold cross-validated within the dataset by using a portion of the experimental trials to estimate calibration coefficients and measuring the resulting endpoint error in the remaining trials. Data were partitioned in 11 subsamples of 9 trials; endpoint errors were calculated for each subsample using calibration coefficients computed from the other 10 subsamples. As we calibrated gaze samples using the same locations that served as targets and as we used the entire dataset, subjects' endpoint errors described in the result section are equivalent to their calibration error.

Eye blinks were detected and removed according to the method suggested by Mathôt (2013); recording epochs of 50 ms before and after a pupil size change velocity of greater than an empirically defined value (3.5×10^4 arbitrary units) were disregarded.

Saccades were defined by the Eyelink saccade detection algorithm (SR Research, 2017) with default thresholds of 35 °/s velocity and 9500 °/s² acceleration. A subset of the detected saccades in each group was visually inspected to confirm that saccades were correctly detected. The better eye was used for each subject, that is, the eye with the lowest median endpoint error.

Detection of visually guided eye movements

Typical eye movement analysis is based on an alternating pattern of fixation and saccades. By contrast, in individuals with nystagmus the eyes are almost never at rest; fixation periods are difficult to define and voluntary saccades are difficult to differentiate from the fast phase of the nystagmus. In addition, a pattern of alternating slow drifts and saccades was often observed (e.g. Worfolk & Abadi, 1991). The idiosyncratic pattern of these eye movements required a new definition of the start and end of a visually guided eye movement applicable to both typical and atypical eye movements.

Visually guided eye movements can be described by the distance between gaze and targets (Fig. 1 C & D). During a gaze shift from a fixation target to the saccade target, the distance of gaze to the fixation target increases (\bar{F} , black line), while the distance to the saccade target decreases (\bar{T} , red line). Subtracting \bar{F} from \bar{T} produces a singular measure of gaze position relative to the two targets: ΔFT (Fig. 1 E, F & I). As soon as a gaze shift takes place, ΔFT moves from

positive values to negative values. A value of zero ($\Delta FT = 0$) indicates that gaze is equidistant from the fixation target and the saccade target. ($\bar{F} = \bar{T}$). Before this zero point (here further referred to as *midpoint*) gaze is closer to the fixation target and after the midpoint gaze is closer to the saccade target. The midpoint can unambiguously be detected in both typical eye movements and in eye movements superimposed by nystagmus, provided that a gaze shift from the fixation target towards the saccade target took place.

If a midpoint was detected ($\Delta FT = 0$) within 1.5 s after target onset, we defined the startpoint of a visually guided eye movement as the point in time at which gaze had reached 7.5% of the total ΔFT displacement. Correspondingly, the time point when the gaze passed 92.5% of the total ΔFT displacement was used as endpoint measure (see Fig. 1 E, F). When ΔFT never crossed the midpoint, the trial was considered a failed gaze shift. Importantly, the same analysis approach was used for all subjects in the present study.

The pattern of results reported in the result section did not considerably change with cut-off values ranging from 1% to 10% in steps of 0.5% (see Fig. A-5 to A-7 in the Appendix). The cut-off values of 7.5% and 92.5% of the ΔFT displacement were chosen based on a validation done in the SC group. For SC individuals, start and endpoint classification as defined by the described new method was compared to the values provided by the eye-tracker parser at the different cut-off values. Normally sighted individuals typically reached the targets within two saccades. Thus, in order to be able to directly compare the new and the eye tracker derived parameters, the endpoint of the gaze shift was defined as the end of the second saccade, if more than one saccade were detected. Next we computed the difference between the estimates of the new and eye tracker based methods. The smallest difference was observed for a cut-off value of 7.5%/92.5%. The new method estimated startpoints as on average 6.8 ms later (range of individual subjects 6-8 ms) and endpoints as on average 7 ms earlier (range for individual subjects: 0-18 ms). These results provide a convincing evaluation of our new method and demonstrate that gaze start- and endpoints were reliably evaluated by the newly introduced method.

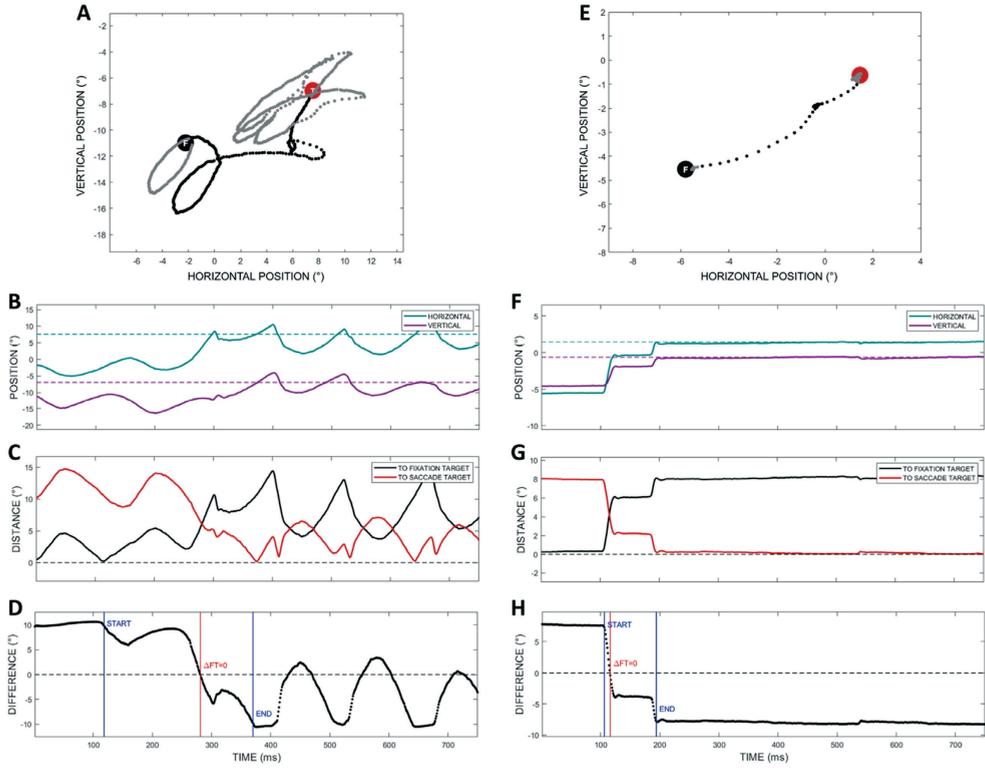


Figure 1: Illustration of the eye movement classification algorithm on two example trials. Left (A - D): single trial of eye movement data from a congenital cataract reversal individual. Right (E - H): single trial of eye movement data from a normal sighted control. A) & E) show 2D plots of the eye movement traces. Black dots indicate samples classified as the visually target guided portion of the eye movement trace while grey dots represent other samples. B) & F) display a horizontal (cyan line) and vertical (magenta line) eye movement trace. The dashed cyan and magenta lines indicate the horizontal and vertical target position. Position 0° represents the centre of the screen. C) & G) show the distance of gaze from fixation target (black line) and target position (red line), respectively. D) & H) the black lines indicate the difference between the distance of gaze from fixation target and saccade target. $\Delta FT = 0$ represents the eye movement midpoint, where gaze is equidistant from the fixation target and saccade target, positive values indicate gaze being closer to the fixation target and negative values indicate gaze being closer to the saccade target position.

Visually guided eye movements measures

Successful eye movements

We considered an eye movement as successful when the gaze position was shifted toward the target within the first 1.5 s of its onset (i.e., the trial contained a midpoint), irrespectively of whether the gaze shift was achieved by one or multiples drifts, saccades, or a combination of both. In this case a midpoint was defined while else the trial was considered a failed visually guided eye movement. To quantify whether a successful eye movement contained a saccade we examined whether the SR saccade detection algorithm had detected a saccade of at least 2° in amplitude within 200 ms before and 200 ms after the midpoint. We assessed the proportions of successful eye movement trials and the proportion of successful trials containing a saccade. The remaining analyses were performed only for successful eye movement trials. Saccade velocity, amplitude, and latency were only assessed for saccades with peak velocities below 1000 °/s (percentage of trials discarded from successful trials containing saccades: CC group: 4 %, subject range: 0 - 9; NC group: 2 %, subject range: 0 - 10; DC group: 1 %, subject range: 0 - 5; SC group: 1 %, subject range: 0 - 1). The cut-off value of 1000 °/s was chosen slightly above normal saccade peak velocities, which in humans typically reach up to 900 °/s (Bahill et al., 1975) to allow for the inclusion of possible abnormally fast saccades in individuals with nystagmus. For each analysis of the remaining eye movement parameters, we report the percentage of discarded trials relative to the number of trials containing a midpoint (successful trials).

Latency and duration of visually guided eye movements

The latency of a visually guided eye movement was defined as the time interval between target onset and the start of a gaze shift (see above). Trials with latencies below 50 ms and above 1000 ms were not analysed for latency and duration (trials discarded: CC group: 22 % (subject range: 8 - 33), NC group: 21 % (subject range: 13 - 42), DC group: 6 % (subject range: 0 - 20), SC group: 3 % (subject range: 0 - 9)). The relatively low cut-off value of 50 ms was chosen just below the minimal latency of typical express saccades in humans (around 75-85 ms; Fischer & Ramsperger, 1984; Biscaldi et al., 1996) in order to include potentially fast latencies due to an atypical release of cortical inhibition as expected for CC individuals. The upper cut-off value of 1000 ms was applied to exclude gaze shifts that began long after target

onset and very likely reflect lapses of concentration. The latency of visually guided eye movements was calculated with the new detection method. Thus, latencies of gaze shifts initiated with a drift and the latencies of gaze shifts initiated with a saccade are not distinguished. Therefore, we additionally report *saccade latencies*, which were defined exclusively as the time interval between target onset and the beginning of detected saccades.

The duration of a visually guided eye movement was defined as the time interval between the start and the end of a gaze shift. In the present study this duration might comprise drifts, one or multiple saccades or a combination of drifts and saccades. Thus, the eye movement duration does not necessarily reflect the duration of a single drift or saccade, as is usually reported in the eye movement literature. The duration of eye movements as reported in the present study indicates the time-efficiency of eye movements from the fixation to the visual target once they were initiated.

Saccade peak velocity

In order to obtain eye movement velocity, we used a moving window technique: A second order polynomial was fitted to the gaze position data (in this case to seven samples, that is, three to the left and three to the right of the central sample). The parameters of the fitted function allowed us to compute the velocity for the central point of the window (see Smeets & Hooge, 2003). *Saccade peak velocity* describes the largest instantaneous velocity within a saccade. If more than one saccade was detected in a trial, the saccade with the largest amplitude was chosen. *Saccade amplitude* refers to the Euclidean distance between the start and endpoint of a detected saccade.

Statistical analysis

Statistics were computed in R (R Core Team, 2018). To be able to quantify evidence for or against the two alternative hypotheses (difference between groups vs. no difference between groups) a Bayesian approach was adopted. The mean and variance of the mean for each variable in each group was estimated by a random intercept model fitted with the R package *brms* (Bürkner, 2017) and the probabilistic programming language *Stan* (Carpenter et al., 2017), using default *brms* estimation priors. Endpoint error, fixation stability, latency, duration, saccade peak velocity, amplitude, main

sequence slope and intercept were modelled as *Student's t*-distributions with variance and shape as free parameters. Values of eye movement success and saccade success were modelled as binomial distributions. Values of endpoint error, fixation stability, latency and duration were log transformed prior to analysis. Posterior distributions are reported in terms of their mean and 95% *Highest Density Interval* (HDI). Bayesian correlation analyses were performed in JASP (JASP Team, 2019), using default priors.

Hypothesis tests were performed by the R package *bain* (Gu et al., 2019), which computes Bayes factors through the Savage-Dickey ratio by using a fraction of the information in the posterior distribution estimated in *brms* to specify the variance of the prior distribution (O'Hagan, 1995; Gu et al., 2018). Unless otherwise described, hypothesis test results are reported as Bayes factors (*BF*) in favour of the undirected hypothesis of a difference between groups. Bayes factors here indicate how much more probable it is that there is a difference between groups than that there is none, given the observed data. $BF < 1$ indicates that it is more probable that there is no difference between the groups, while $BF > 1$ indicates that it is more probable that there is a difference between groups.

In line with Jeffreys (1961), we evaluate *BF*'s greater than 3 as “evidence worth considering” for a relevant group difference and *BF*'s smaller than 0.33 as evidence for no group difference; we refer to comparisons with a *BF* between 0.33 and 3 as “inconclusive marginal evidence”, and indicate the direction of the trend. $BF > 10$ were rounded to the nearest digit. In addition, *BF*s were reported in $\ln(BF)$, the natural logarithm of the Bayes factor in Table 2.

All data, analysis and stimulus scripts will be made available on reasonable request by the corresponding author.

RESULTS

Success and accuracy of visually guided eye movements

All subjects in all groups were able to perform visually guided eye movements, albeit with considerable differences in eye movement trajectory. Fig. 1 displays two example trials, one from a CC individual (Fig. 1 A - D) and one from a SC individual (Fig. 1 E - H) starting at the time point at which the target jumped from its previous position (the

fixation target) to a new target position (the saccade target). SC individuals and all but one DC individuals consistently made typical rapid eye movements (saccades) to the saccade target (Fig. 3 B & F). In contrast, eye movement trajectories in the CC and NC individuals comprised, as expected, by saccades and slow drifts towards the target. Extensive nystagmus during task performance emerged in all CC individuals but in only 1 of 16 individuals in the DC group.

The eye movement signal can be transformed into two components: displacement parallel to the axis between fixation target and saccade target and displacement orthogonal to this axis. Plots of the parallel (grey solid lines in Fig. 2) and orthogonal (cyan dotted lines in Fig. 2) component traces 200 ms before until 200 ms after the eye movement midpoint (see *Detection of visually guided eye movements* section) are displayed in Fig. 2 A for one example subject from each group and in Fig. 2 B for all subjects from each group. All subjects predominantly made eye movements along the parallel axis indicating that they all were able to systematically perform eye movements from the fixation target to the saccade target, albeit with considerably variable trajectories in the CC and NC groups. The variability in the orthogonal component in the CC and NC groups reflects eye movements unrelated to the intended gaze shift, such as nystagmus.

CC individuals performed successful visually guided eye movements in 84% of the trials (95% HDI: [74, 91], range: 60 - 94%), which was lower compared to both the success rate of the SC group ($BF = 14$; see Table 2 for detailed statistical estimates of this and following measures). The comparison to the DC group was inconclusive, with marginal evidence indicating a lower success rate in the CC group ($BF = 1.66$). Importantly, the success rate of the CC group was, however, similar to that of the NC group ($BF = 0.14$). Of these successful visually guided eye movements, 79 % contained at least one saccade in the CC group (95 % HDI: [68, 88], range: 66 - 91 %; Fig. 3 B), which was a lower rate than found both in the SC group ($BF = 4.2 \times 10^4$) and the DC group ($BF = 95$). The comparison to the NC group was inconclusive, with marginal evidence indicating that the proportion of saccades in successful visually guided eye movements was similar between the CC and the NC group ($BF = 0.44$). Correlations in the CC group between the proportion of successful visually guided eye movements and duration of blindness ($r = 0.24$, $BF = 0.49$) or duration of visual experience since surgery ($r = -0.25$, $BF =$

0.49) were inconclusive with marginal evidence pointing towards an absence of a correlation.

The accuracy of visually guided eye movements was assessed as endpoint error, which represents the position of gaze at the end of an eye movement. The final endpoint error of successful visually guided eye movements in the CC group was 1.44 visual degrees on average (95% HDI: [1.12, 1.84]) and similar to the average endpoint error of the NC group ($BF = 0.24$; see Table 2 for detailed statistical results). The average endpoint error of the CC group was larger than in the SC group ($BF = 4.1 \times 10^4$) and in the DC group ($BF = 913$) (Fig. 3 E). Correlations in the CC group between endpoint error and duration of blindness ($r = 0.33$, $BF = 0.41$) or duration of visual experience since surgery ($r = -0.03$, $BF = 0.56$) were inconclusive with marginal evidence pointing towards an absence of a correlation.

In summary, CC individuals were able to successfully perform visually guided eye movements and achieved an astonishing accuracy. Lower success rates and lower eye movement accuracy in this group could be accounted for by the pathological nystagmus characterizing eye movements in CC individuals.

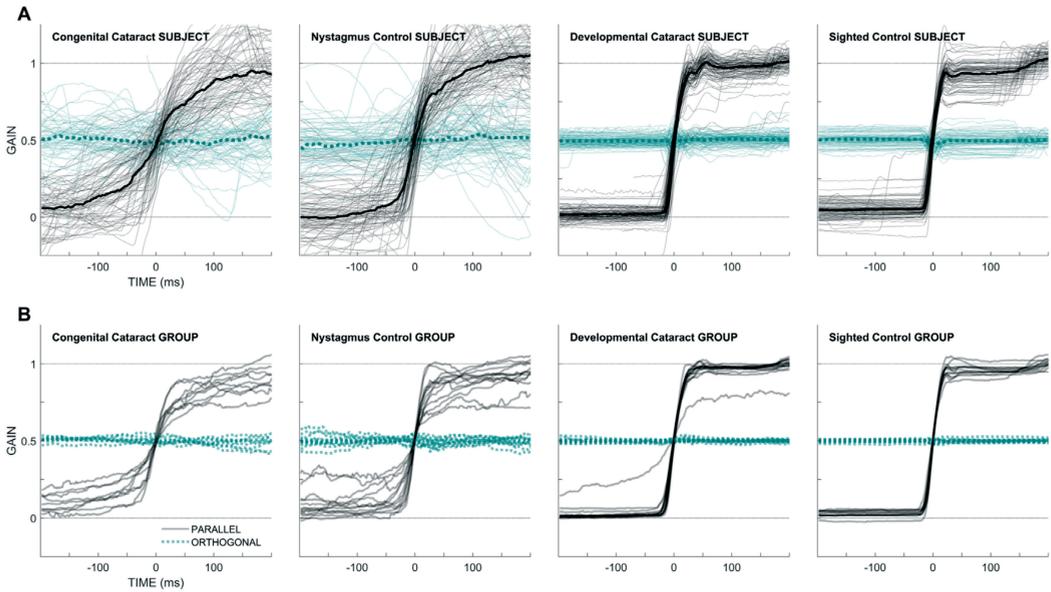


Figure 2. Eye movement traces in the congenital cataract reversal group, the developmental cataract reversal group, a group of participants with nystagmus and in normally sighted controls. Eye movement traces start 200 ms before and end 200 ms after a detected midpoint in a trial, and indicate gaze position after transformation into two components: displacement parallel (grey solid lines) and orthogonal (cyan dotted lines) to the axis between the fixation and the saccade target. Traces were normalized to the distance between the fixation and the saccade target, which are marked with dotted lines at positions 0 and 1, respectively. A) Examples of data recorded in one individual subject of each group. Individual plots show position traces for all trials with a detectable midpoint (thin lines) and the corresponding median trace of a participant (median position at each time point; thick lines). Individual subject plots for each participant can be found in the Appendix. B) Median eye position traces for each subject and component per group.

Latency, duration and velocity of visually guided movements

Next we evaluated the timing (initiation time, duration and velocity) of visually guided eye movements in CC individuals. The latency to initiate an eye movement indicates the processing time required to disengage from the current fixation target, program an eye movement and initiate the motor commands. The duration of a visually guided eye movement evaluates the time-efficiency of movement execution. For example, the duration of eye movements is expected to be longer, if less efficient movements, such as drifts, emerge or when pathological eye movements, such as nystagmus, superimpose the visually guided eye movement. Finally, the peak velocity of saccades (when present) was compared across groups.

In the CC group, the latency to initiate a visually guided eye movement was on average 220 ms (95% HDI: [199, 244], see Fig.3 C and Table 2), and was similar to the average latencies in the NC group ($BF = 0.15$) and the SC group ($BF = 0.17$). A comparison of latencies between the CC and DC groups was inconclusive ($BF = 1.05$). The latencies analysed here comprised both visually guided eye movements starting with a saccade and visually guided eye movements starting with a drift. For a better comparison to previous studies (Dunn et al., 2015; Huurneman et al., 2016; Worfolk & Abadi, 1991) we additionally computed the latency to the first saccade within a visually guided movement (when present). In the CC group, the latency to initiate a visually guided saccade was on average 298 ms (95% HDI: [267, 333], see Table 2 for all groups estimates), which was longer than the average saccade latency in the SC group ($BF = 4.6 \times 10^4$). The comparison of the CC group with the DC group ($BF = 1.59$) was inconclusive with marginal evidence indicating lower saccade latencies in the CC group. Saccade latencies were similar between the CC and NC group ($BF = 0.33$).

The duration of visually guided eye movements was longer in the CC group than in the SC group ($BF = 4.0 \times 10^3$) and in the DC group ($BF = 1.0 \times 10^3$) but was similar between the CC and the NC groups ($BF = 0.19$) (Fig. 3 D).

Saccade peak velocities in the CC group were lower than in all other groups (Fig. 3 G), that is, lower than in the SC group ($BF = 2.5 \times 10^5$), the DC group ($BF = 48$) and the NC group ($BF = 302$). As saccade peak velocity depends on the amplitude of a saccade, lower saccade peak velocities might be a consequence of hypometric

saccades in CC individuals. In fact, saccade amplitude was lower in the CC group than both in the SC group ($BF = 217$) and in the DC group ($BF = 630$) (Fig. 3 H). The comparison of average saccade amplitudes between the CC group and the NC group was inconclusive, with marginal evidence indicating no difference between these groups ($BF = 0.6$). To further clarify this result, we analysed the saccade main sequence (Bahill, 1975). As in log-log space saccade amplitude and saccade peak velocities are typically linearly related, we fitted linear regression lines relating saccade amplitude and peak velocity for each subject in the CC and NC groups. The mean of the slope coefficients was similar in the CC and NC groups ($BF = 0.23$), but the mean of the intercept coefficients was lower in the CC than in the NC group ($BF = 39$). Thus, lower saccade velocities in the CC group cannot be accounted for by lower saccade amplitudes overall, as peak velocities were lower for the whole range of amplitudes (see also Fig. A-9 in the Appendix).

In summary, CC individuals were able to initiate and complete eye movements to visual targets as fast as NC and SC controls. By contrast, the peak velocity of saccades (when present) in the CC group was lower than in all other groups, including the NC group.

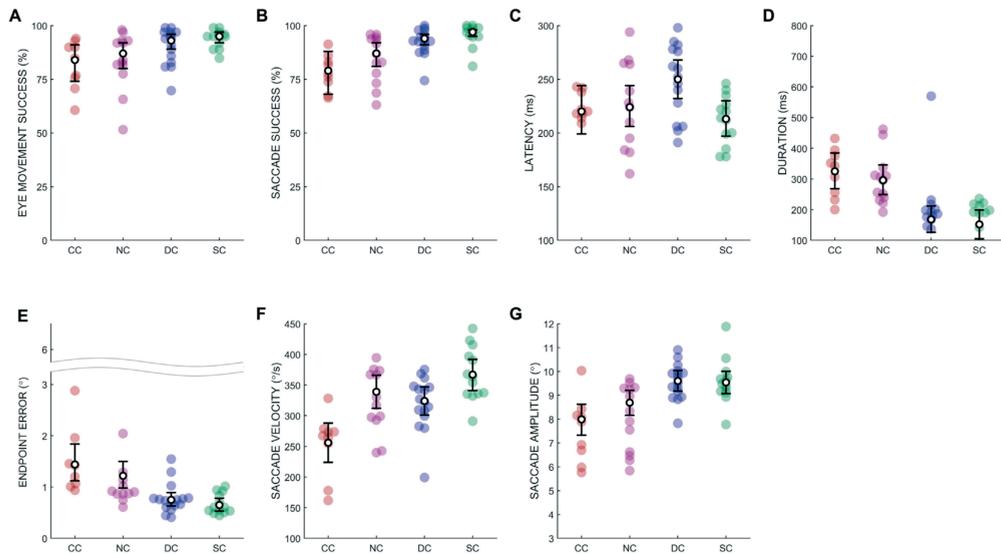


Figure 3. Eye movement characteristics for all groups. Coloured, filled circles represent individual subject medians; black, open circles represent estimated posterior distribution means per group, and bars represent 95% highest density intervals (HDI). CC = congenital cataract group; NC = nyctagmus control group; DC = developmental cataract group; SC = normally sighted control group. A) Percentage of successful visually guided eye movements. B) Percentage of successful eye movements containing a saccade. C) Eye movement latencies. D) Eye movement durations. E) Endpoint errors. F) Saccade peak velocities. G) Saccade amplitudes.

Parameter (unit)	Group	Mean estimate	95% HDI	BF	ln(BF)
Eye movement success, %	CC	0.84	(0.74, 0.91)	—	—
	NC	0.87	(0.8, 0.92)	0.17	-1.772
	DC	0.93	(0.89, 0.96)	1.66	0.507
	SC	0.95	(0.92, 0.97)	14	2.629
Saccade success, %	CC	0.79	(0.68, 0.88)	—	—
	NC	0.87	(0.81, 0.92)	0.44	-0.821
	DC	0.94	(0.91, 0.96)	95	4.556
	SC	0.97	(0.95, 0.98)	4.2×10^4	10.645
Latency, ms	CC	220	(199, 244)	—	—
	NC	224	(206, 244)	0.15	-1.897
	DC	250	(232, 268)	1.05	0.049
	SC	213	(197, 230)	0.17	-1.772
Saccade latency, ms	CC	298	(267, 333)	—	—
	NC	272	(248, 298)	0.33	-1.109
	DC	256	(236, 277)	1.59	0.464
	SC	208	(189, 226)	4.6×10^4	10.728
Duration, ms	CC	325	(268, 385)	—	—
	NC	296	(249, 346)	0.19	-1.661
	DC	168	(126, 212)	1.0×10^3	6.914
	SC	152	(105, 199)	4.0×10^3	8.302
Endpoint error, °	CC	1.44	(1.12, 1.84)	—	—
	NC	1.22	(0.98, 1.5)	0.24	-1.427
	DC	0.75	(0.63, 0.89)	913	6.817
	SC	0.65	(0.53, 0.78)	4.1×10^4	10.628
Saccade peak velocity, °/s	CC	256	(224, 288)	—	—
	NC	339	(312, 366)	302	5.71
	DC	324	(301, 347)	48	3.865
	SC	367	(341, 392)	2.5×10^5	12.409
Saccade amplitude, °	CC	7.99	(7.33, 8.62)	—	—
	NC	8.69	(8.16, 9.21)	0.6	-0.511
	DC	9.6	(9.17, 10.04)	630	6.445
	SC	9.54	(9.07, 10.01)	217	5.378

Table 2. Statistical results and parameter estimates. *Notes:* *BFs* indicate evidence in favor of a difference between groups relative to no difference between groups. *BFs* reflect comparisons of the CC group with each of the other three groups. HDI represents the 95% highest density interval of the posterior distribution of group means.

DISCUSSION

In the present study we investigated the presence and characteristics of visually guided eye movements in individuals with a history of congenital, dense, total, bilateral cataracts (CC group) which were later removed at different times in their lives. Visually guided eye movements in CC individuals were compared to those of normally sighted controls (SC group) and to eye movements of a group of developmental cataract individuals (DC group). An additional group of individuals who suffered nystagmus but who did not have a history of congenital loss of pattern vision was tested as an additional control group (NC group) to control for the nystagmus prevalent in all

individuals with a history of congenital cataracts. All CC individuals were able to execute visually guided eye movements. Except for a lower saccade velocity, differences between the CC and the SC groups in eye movement success, latency, duration, endpoint error, and saccade amplitude could be accounted for by the nystagmus superimposing visually guided eye movements. To the best of our knowledge, the present data provide the first quantitative description of visually guided eye movements following long-lasting visual deprivation (more than three years) from birth and demonstrate a remarkable ability to perform visually guided movements in this group. This result contrasts with the often extensive and prevailing visual deficits in CC individuals. These findings suggest that visual impairments are likely not predominantly linked to aberrant basic visually guided eye movements.

Commonly used calibration and analysis routines designed for normally sighted individuals are not well suited for individuals with nystagmus. In infantile nystagmus, drift and saccades alternate in a highly heterogeneous fashion and often vary extensively both within and between individuals. Recently a calibration method has been suggested which was based on the automatic extraction of foveation periods according to relative velocity within the nystagmus slow phase periods, outlier correction, and waveform shape comparison (Dunn et al., 2019; Rosengren et al., 2019). In the present study we propose a different approach, which defines the eye movements' start- and endpoints independent of the shape and the uniformity of nystagmus waveforms. Based on this definition and using a cross-validation approach, calibration coefficients were iteratively optimized to minimize the endpoint error over the full set of experimental data, rather than relying on a short, initial calibration procedure. Our method resulted in a mean deviation of endpoint error estimates of no more than one degree of visual angle. Moreover, we successfully adjusted and validated the new method in the SC group by comparing the classification of start- and endpoints of the new and the standard method offered by the eye tracker software.

Previous studies have not systematically assessed individual eye movement characteristics of visually guided eye movements in CC individuals. Some qualitative reports described gaze following behavior in a child after cataract surgery (Chen et al., 2016). Moreover, clinical studies investigating eye movements in CC individuals have exclusively focused on a description of nystagmus waveform characteristics (Abadi, 2002; Abadi et al., 2006; Birch et al., 2012),

without assessing the presence or quantitative characteristics of visually guided eye movements. For NC individuals it has been shown that nystagmus does not prevent visually guided eye movements (e.g. Yee et al., 1976; Kommerell, 1986; Collewijn et al., 1985). However, only a few studies have reported quantitative metrics such as saccade latency and amplitude (Worfolk & Abadi, 1991; Dunn et al., 2015; Huurneman et al., 2016) in individuals with nystagmus. Two of these studies found longer saccade latencies in NC individuals compared to healthy controls (Dunn et al., 2015; Huurneman et al., 2016). The present study replicated this finding for individuals of the NC group and extended it to CC individuals. Importantly, it has been shown that individuals with infantile nystagmus seem to be able to strategically adapt their nystagmus slow and fast phases, using saccades when targets were located in the direction of the nystagmus quick phase, and using drifts when the target was located in the opposite direction (Kommerell, 1986; Worfolk & Abadi, 1991). This pattern of eye movements implies that drifts need to be considered as a crucial element of visually guided eye movements in individuals with infantile nystagmus. Therefore, we suggest that an estimate of latency must take into account any type of eye movement in the direction of the target in order to be a valid estimate of eye movement initiation time for individuals with nystagmus. With such a latency measure we did neither find a difference between the NC and the SC group, nor did we find evidence for a difference between the CC and the SC group, suggesting that visually guided eye movements in CC individuals were programmed as fast as in normally sighted individuals despite their nystagmus.

We demonstrated further that visually guided eye movements in CC individuals were remarkably accurate, despite their nystagmus and their overall low visual acuity. The endpoint error of the CC group was indistinguishable from the endpoint error the NC group, and on average only 0.79° larger than in normally sighted controls. This present observation in CC individuals extends a previous report indicating that individuals with nystagmus can accurately fixate visual targets (Dell'Osso et al., 1992).

CC individuals were investigated between 1 and 19 years after surgery. In order to maximize the likelihood that CC individuals indeed suffered a complete loss of pattern vision at birth we employed multiple inclusion criteria, which were only met by a subgroup of the congenital cataract individuals treated at the LV Prasad Eye Institute. These criteria included family history and reports, lack of fundus

visibility, strabismus, the presence of nystagmus, partially absorbed cataracts and the results of the pre-surgery vision assessment. Moreover, all CC individuals still suffered marked visual impairments after cataract removal surgery. A limitation of studies in congenital cataract reversal individuals is, that while highly likely, there is no absolute guarantee that all participants had dense and total cataracts at birth. Thus, the most conservative conclusion would be, that visually guided eye movements depend less on good visual capacities at birth than a number of other visual functions (Lewis & Maurer, 2005; Ganesh et al., 2014).

We can only speculate on how CC individuals achieve the impressive ability to execute visually guided eye movements. Multiple, parallel, subcortical and cortical pathways are involved in oculomotor control (Pierrot-Deseilligny et al., 2004; Lynch & Tian, 2006). Thus, visually guided eye movements of CC individuals could be mediated by an experience independent development or a recovery of either subcortical and/or cortical pathways. Overt orienting to visual stimuli as measured in the present study can be subcortically generated, that is, through the retino-collicular pathway. In fact, the subcortical pathway has for a long time been considered to be sufficient for the generation of visually guided saccades, since non-human animal studies have observed visually guided eye movements even after damaging occipital, parietal and frontal cortex (Sprague, 1966; Humphrey, 1974; Schiller et al., 1980; Trevathe, 1968; Schneider, 1969). Moreover, the retino-cortical pathway of primates is known to be highly mature at birth (Stein, 1984; Qu et al., 2016; Wallace et al., 1997). It has been hypothesized that in human infants, during the first months of live visually guided orienting is largely subcortically mediated. By contrast, cortical systems have a more protracted development and are thought to start guiding visual behavior only later (Bronson, 1974; Johnson, 1990). In human adults, only lesions in the parietal eye field or of parietal projections to the superior colliculus have consistently been associated with deficits in the production of visually guided eye movements (Pierrot-Deseilligny et al., 1991; Gaymard et al., 2003; Müri & Nyffeler, 2008). Electrophysiological recordings in visually deprived monkeys' Brodmann area 7 did not find a considerable recovery of visual responses at the end and one year after the end of the visual deprivation (Hyvärinen et al., 1981; Carlson et al., 1987). Area 7 in monkeys is a multisensory parietal region adjacent to the LIP area, which is thought to be the homologue of the human parietal eye field. Therefore, we speculate that the CC individuals' visually guided eye movements might be predominantly

mediated by retino-collicular mechanisms. In fact, other tasks that have been associated with the superior colliculus, such as multisensory redundancy effects in simple target detection tasks have been found to be unimpaired in CC individuals (Putzar et al., 2012; de Heering et al., 2016). Future studies must implement more complex active vision tasks in order to isolate the functionality of cortical eye movement related pathways.

Despite the finding of unimpaired (compared the NC individuals) major parameters of visually guided eye movements, the peak velocity of saccades was found to be lower in the CC group than in the NC group, that is, the lower velocity could not be explained by nystagmus. It has been demonstrated that in accord with the dual coding hypothesis (Sparks & Mays, 1990) both saccade trajectories and eye movement kinematics such as velocity are coded by the superior colliculus (Goossens & van Opstal, 2006, 2012; Smalianchuk et al., 2018). Studies in visually deprived cats observed a lower number of visually responsive cells and a lower response rate of visual neurons (Rauschecker & Harris, 1983). Thus, it might be speculated that the lower velocity of the visually guided eye movements of CC individuals might be related to changes in the response rate of visual collicular neurons, while the high accuracy of their visually guided eye movements might be related to the activation of a preserved topological organization of the visually responsive neurons. In fact, non-human animal studies have found a typical topography and typical receptive field sizes in the superior colliculus of visually deprived animals at the end of the deprivation period (Vidyasagar, 1978; King & Carlile, 1993). In accord with this finding a recent study found a retinotopic organization and typical latency of the first visual response (Sourav, 2018).

In summary, the present study demonstrated a remarkable ability of cataract reversal individuals with a history of long-lasting loss of pattern vision from birth to execute visually guided eye movements despite suffering nystagmus.

REFERENCES

- Abadi, R V. (2002). Motor and sensory characteristics of infantile nystagmus. *British Journal of Ophthalmology*, 86(10), 1152–1160. <https://doi.org/10.1136/bjo.86.10.1152>
- Abadi, Richard V., Forster, J. E., & Lloyd, I. C. (2006). Ocular motor outcomes after bilateral and unilateral infantile cataracts. *Vision Research*, 46(6–7), 940–952. <https://doi.org/10.1016/j.visres.2005.09.039>
- Alahyane, N., Lemoine- Lardennois, C., Tailhefer, C., Collins, T., Fagard, J., & Doré-Mazars, K. (2016). Development and learning of saccadic eye movements in 7- to 42-month-old children. *Journal of Vision*, 16(1), 6. <https://doi.org/10.1167/16.1.6>
- Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception & Psychophysics*, 17(3), 293–302. <https://doi.org/10.3758/BF03203214>
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical biosciences*, 24(3-4), 191-204.
- Birch, E. E., Wang, J., Felius, J., Stager, D. R., & Hertle, R. W. (2012). Fixation control and eye alignment in children treated for dense congenital or developmental cataracts. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 16(2), 156–160. <https://doi.org/10.1016/j.jaapos.2011.11.007>
- Biscaldi, M., Fischer, B., & Stuhr, V. (1996). Human express saccade makers are impaired at suppressing visually evoked saccades. *Journal of Neurophysiology*, 76(1), 199-214.
- Bottari, D., Troje, N. F., Ley, P., Hense, M., Kekunnaya, R., & Röder, B. (2015). The neural development of the biological motion processing system does not rely on early visual input. *Cortex*, 71, 359-367.
- Bottari, D., Kekunnaya, R., Hense, M., Troje, N. F., Sourav, S., & Röder, B. (2018). Motion processing after sight restoration: No competition between visual recovery and auditory compensation. *NeuroImage*, 167, 284–296. <https://doi.org/10.1016/j.neuroimage.2017.11.050>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Brenner, E., Cornelissen, F., & Nuboer, W. (1990). Striking absence of long-lasting effects of early color deprivation on monkey vision. *Developmental Psychobiology*, 23(5), 441–448. <https://doi.org/10.1002/dev.420230506>
- Bronson, G. (1974). The Postnatal Growth of Visual Capacity. *Child Development*, 45(4), 873–890.
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- Carlson, S., Pertovaara, A., & Tanila, H. (1987). Late effects of early binocular visual deprivation on the function of Brodmann's area 7 of monkeys (*Macaca arctoides*). *Developmental Brain Research*, 33(1), 101–111. [https://doi.org/10.1016/0165-3806\(87\)90180-5](https://doi.org/10.1016/0165-3806(87)90180-5)
- Carlson, S. (1990). Visually Guided Behavior of Monkeys After Early Binocular Visual Deprivation. *International Journal of Neuroscience*, 50(3–4), 185–194. <https://doi.org/10.3109/002074590008987171>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P(M), & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).

- Carrasco, M. M., Razak, K. A., & Pallas, S. L. (2005). Visual Experience Is Necessary for Maintenance But Not Development of Receptive Fields in Superior Colliculus. *Journal of Neurophysiology*, 94(3), 1962–1970. <https://doi.org/10.1152/jn.00166.2005>
- Chen, J., Wu, E. D., Chen, X., Zhu, L. H., Li, X., Thorn, F., Ostrovsky, Y., & Qu, J. (2016). Rapid integration of tactile and visual information by a newly sighted child. *Current Biology*, 26(8), 1069–1074.
- Cohen, M. E., & Ross, L. E. (1978). Latency and accuracy characteristics of saccades and corrective saccades in children and adults. *Journal of Experimental Child Psychology*, 26(3), 517–527. [https://doi.org/10.1016/0022-0965\(78\)90130-3](https://doi.org/10.1016/0022-0965(78)90130-3)
- Collewijn, H., Apkarian, P., & Spekreijse, H. (1985). The oculomotor behaviour of human albinos. *Brain*, 108(1), 1–28. <https://doi.org/10.1093/brain/108.1.1>
- de Heering, A., & Maurer, D. (2014). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology*, 56(1), 96–108.
- de Heering, A., Dormal, G., Pelland, M., Lewis, T., Maurer, D., & Collignon, O. (2016). A brief period of postnatal visual deprivation alters the balance between auditory and visual attention. *Current Biology*, 26(22), 3101–3105.
- Dell'Osso, L. F., Van der Steen, J., Steinman, R. M., & Colewijn, H. (1992). Foveation dynamics in congenital nystagmus I: Fixation. *Documenta Ophthalmologica*, 79(1), 1–23.
- Dunn, M. J., Margrain, T. H., Woodhouse, J. M., & Erichsen, J. T. (2015). Visual Processing in Infantile Nystagmus Is Not Slow. *Investigative Ophthalmology & Visual Science*, 56(9), 5094. <https://doi.org/10.1167/iovs.15-16977>
- Dunn, M. J., Harris, C. M., Ennis, F. A., Margrain, T. H., Woodhouse, J. M., McIlreavy, L., & Erichsen, J. T. (2019). An automated segmentation approach to calibrating infantile nystagmus waveforms. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-018-1178-5>
- Elleberg, D., Lewis, T. L., Maurer, D., Hong Lui, C., & Brent, H. P. (1999). Spatial and temporal vision in patients treated for bilateral congenital cataracts. *Vision Research*, 39(20), 3480–3489. [https://doi.org/10.1016/S0042-6989\(99\)00078-4](https://doi.org/10.1016/S0042-6989(99)00078-4)
- Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, 57(1), 191–195.
- Ganesh, S., Arora, P., Sethi, S., Gandhi, T. K., Kalia, A., Chatterjee, G., & Sinha, P. (2014). Results of late surgical intervention in children with early-onset bilateral cataracts. *British Journal of Ophthalmology*, 98(10), 1424–1428.
- Gaymard, B., Lynch, J., Ploner, C. J., Condy, C., & Rivaud-Péchoix, S. (2003). The parieto-collicular pathway: Anatomical location and contribution to saccade generation: Parieto-collicular pathway. *European Journal of Neuroscience*, 17(7), 1518–1526. <https://doi.org/10.1046/j.1460-9568.2003.02570.x>
- Goossens, H. H. L. M., & Van Opstal, A. J. (2006). Dynamic ensemble coding of saccades in the monkey superior colliculus. *Journal of neurophysiology*, 95(4), 2326–2341.
- Goossens, H. H. L. M., & van Opstal, A. J. (2012). Optimal control of saccades by spatial-temporal activity patterns in the monkey superior colliculus. *PLoS Computational Biology*, 8(5), e1002508.
- Gu, X., Mulder, J., and Hoijsink, H. (2018). Approximated adjusted fractional Bayes factors: A general method for testing informative hypotheses. *British Journal of Mathematical and Statistical Psychology*. <http://dx.doi.org/10.1111/bmsp.12110>
- Gu, X., Hoijsink, H., Mulder, J., & van Lissa, C. J. (2019). *bain: Bayes Factors for Informative Hypotheses*. R package version 0.2.1. <https://CRAN.R-project.org/package=bain>

- Hadad, B.-S., Maurer, D., & Lewis, T. L. (2012). Sparing of sensitivity to biological motion but not of global motion after early visual deprivation: Sparing of biological motion after visual deprivation. *Developmental Science*, *15*(4), 474–481. <https://doi.org/10.1111/j.1467-7687.2012.01145.x>
- Hall, E. C., Gordon, J., Hainline, L., Abramov, I., & Engber, K. (2000). Childhood Visual Experience Affects Adult Voluntary Ocular Motor Control: *Optometry and Vision Science*, *77*(10), 511–523. <https://doi.org/10.1097/00006324-200010000-00005>
- Harris, C. M., Jacobs, M., Shawkat, F., & Taylor, D. (1993). The development of saccadic accuracy in the first seven months. *Clinical vision sciences*, *8*(1), 85–96.
- Harwerth, R. S., Smith, E. L., Paul, A. D., Crawford, M. L., & Von Noorden, G. K. (1991). Functional effects of bilateral form deprivation in monkeys. *Investigative ophthalmology & visual science*, *32*(8), 2311–2327.
- Humphrey, N. K. (1974). Vision in a Monkey without Striate Cortex: A Case Study. *Perception*, *3*(3), 241–255. <https://doi.org/10.1068/p030241>
- Huurneman, B., Boonstra, F. N., & Goossens, J. (2016). Perceptual Learning in Children With Infantile Nystagmus: Effects on 2D Oculomotor Behavior. *Investigative Ophthalmology & Visual Science*, *57*(10), 4229. <https://doi.org/10.1167/iov.16-19555>
- Hyvärinen, J., Hyvärinen, L., & Linnankoski, I. (1981). Modification of parietal association cortex and functional blindness after binocular deprivation in young monkeys. *Experimental Brain Research*, *42*(1). <https://doi.org/10.1007/BF00235723>
- Irving, E. L., González, E. G., Lillakas, L., Wareham, J., & McCarthy, T. (2011). Effect of Stimulus Type on the Eye Movements of Children. *Investigative Ophthalmology & Visual Science*, *52*(2), 658. <https://doi.org/10.1167/iov.10-5480>
- JASP Team (2019). JASP (Version 0.11.1)[Computer software]. Retrieved from <https://jasp-stats.org/>
- Jeffreys, H. 1961. *Theory of probability (3rd ed.)*. New York: Oxford University Press.
- Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience*, *2*(2), 81–95.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*(10), 766–774. <https://doi.org/10.1038/nrn1766>
- Kalia, A., Lesmes, L. A., Dorr, M., Gandhi, T., Chatterjee, G., Ganesh, S., Bex, J. P., & Sinha, P. (2014). Development of pattern vision following early and extended blindness. *Proceedings of the National Academy of Sciences*, *111*(5), 2035–2039.
- Kämpf, D., & Piper, H.-F. (1987). Eye movements and vestibulo-ocular reflex in the blind. *Journal of Neurology*, *234*(5), 337–341. <https://doi.org/10.1007/BF00314291>
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.
- Kommerell, G. (1986). Congenital nystagmus: Control of slow tracking movements by target offset from the fovea. *Graefes's Archive for Clinical and Experimental Ophthalmology*, *224*(3), 295–298. <https://doi.org/10.1007/BF02143074>
- King, A. J., & Carlile, S. (1993). Changes induced in the representation of auditory space in the superior colliculus by rearing ferrets with binocular eyelid suture. *Experimental Brain Research*, *94*(3), 444–455.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *410*(6831), 890–890. <https://doi.org/10.1038/35073749>
- Leigh, R. J., & Zee, D. S. (1980). Eye movements of the blind. *Investigative Ophthalmology & Visual Science*, *19*(3), 328–331.

- Lewis, T. L., & Maurer, D. (2005). Multiple sensitive periods in human visual development: evidence from visually deprived children. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 46(3), 163-183.
- Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain and Cognition*, 68(3), 293-308. <https://doi.org/10.1016/j.bandc.2008.08.019>
- Lynch, J., & Tian, J. R. (2006). Cortico-cortical networks and cortico-subcortical loops for the higher control of eye movements. *Progress in Brain Research*, 151, 461-501. [https://doi.org/10.1016/S0079-6123\(05\)51015-X](https://doi.org/10.1016/S0079-6123(05)51015-X)
- Mathôt, S. (2013). A Simple Way to Reconstruct Pupil Size During Eye Blinks. <https://doi.org/10.6084/m9.figshare.688001>
- McKyton, A., Ben-Zion, I., Doron, R., & Zohary, E. (2015). The limits of shape recognition following late emergence from blindness. *Current Biology*, 25(18), 2373-2378.
- Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research*, 121(4), 391-400. <https://doi.org/10.1007/s002210050473>
- Müri, R. M., & Nyffeler, T. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades as revealed by lesion studies with neurological patients and transcranial magnetic stimulation (TMS). *Brain and Cognition*, 68(3), 284-292. <https://doi.org/10.1016/j.bandc.2008.08.018>
- O'Hagan, A. (1995). Fractional Bayes factors for model comparison (with discussion). *Journal of the Royal Statistical Society. Series B*, 57, 99-138.
- Pierrot-Deseilligny, Ch., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain*, 114(3), 1473-1485. <https://doi.org/10.1093/brain/114.3.1473>
- Pierrot-Deseilligny, Charles, Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex: *Current Opinion in Neurology*, 17(1), 17-25. <https://doi.org/10.1097/00019052-200402000-00005>
- Pitchaimuthu, K., Sourav, S., Bottari, D., Banerjee, S., Shareef, I., Kekunnaya, R., & Röder, B. (2019). Color vision in sight recovery individuals. *Restorative Neurology and Neuroscience*, 37(6), 583-590.
- Putzar, L., Hötting, K., Rösler, F., & Röder, B. (2007). The development of visual feature binding processes after visual deprivation in early infancy. *Vision research*, 47(20), 2616-2626.
- Putzar, L., Hötting, K., & Röder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restorative neurology and neuroscience*, 28(2), 251-257.
- Putzar, L., Gondan, M., & Röder, B. (2012). Basic multisensory functions can be acquired after congenital visual pattern deprivation in humans. *Developmental neuropsychology*, 37(8), 697-711.
- Qu, J., Zhou, X., Zhu, H., Cheng, G., W.S. Ashwell, K., & Lu, F. (2006). Development of the human superior colliculus and the retinocollicular projection. *Experimental Eye Research*, 82(2), 300-310. <https://doi.org/10.1016/j.exer.2005.07.002>
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- Rauschecker, J. P., & Harris, L. R. (1983). Auditory compensation of the effects of visual deprivation in the cat's superior colliculus. *Experimental Brain Research*, 50(1), 69-83.

- Regal, D. M., Boothe, R., Teller, D. Y., & Sackett, G. P. (1976). Visual acuity and visual responsiveness in dark-reared monkeys (*Macaca nemestrina*). *Vision research*, *16*(5), 523-530.
- Röder, B., Ley, P., Shenoy, B. H., Kekunnaya, R., & Bottari, D. (2013). Sensitive periods for the functional specialization of the neural system for human face processing. *Proceedings of the National Academy of Sciences*, *110*(42), 16760-16765.
<https://doi.org/10.1073/pnas.1309963110>
- Rosengren, W., Nyström, M., Hammar, B., & Stridh, M. (2019). A robust method for calibration of eye tracking data recorded during nystagmus. *Behavior Research Methods*.
<https://doi.org/10.3758/s13428-019-01199-0>
- Ross, R. G., Radant, A. D., Hommer, D. W., & Young, D. A. (1994). Saccadic eye movements in normal children from 8 to 15 years of age: A developmental study of visuospatial attention. *Journal of Autism and Developmental Disorders*, *24*(4), 413-431.
<https://doi.org/10.1007/BF02172126>
- Schiller, P. H., True, S. D., & Conway, J. L. (1980). Deficits in eye movements following frontal eye-field and superior colliculus ablations. *Journal of Neurophysiology*, *44*(6), 1175-1189. <https://doi.org/10.1152/jn.1980.44.6.1175>
- Schneider, G. E. (1969). Two Visual Systems. *Science*, *163*(3870), 895-902.
<https://doi.org/10.1126/science.163.3870.895>
- Schneider, R. M., Thurtell, M. J., Eisele, S., Lincoff, N., Bala, E., & Leigh, R. J. (2013). Neurological Basis for Eye Movements of the Blind. *PLoS ONE*, *8*(2), e56556.
<https://doi.org/10.1371/journal.pone.0056556>
- Schulze-Bonsel, K., Feltgen, N., Burau, H., Hansen, L., & Bach, M. (2006). Visual acuities "hand motion" and "counting fingers" can be quantified with the Freiburg visual acuity test. *Investigative Ophthalmology & Visual Science*, *47*(3), 1236-1240.
- Smalianchuk, I., Jagadisan, U. K., & Gandhi, N. J. (2018). Instantaneous Midbrain Control of Saccade Velocity. *The Journal of Neuroscience*, *38*(47), 10156-10167.
<https://doi.org/10.1523/JNEUROSCI.0962-18.2018>
- Smeets, J. B., & Hooge, I. T. (2003). Nature of variability in saccades. *Journal of neurophysiology*, *90*(1), 12-20.
- Sourav, S., Bottari, D., Kekunnaya, R., & Röder, B. (2018). Evidence of a retinotopic organization of early visual cortex but impaired extrastriate processing in sight recovery individuals. *Journal of vision*, *18*(3), 22-22.
- Sparks, D. L., & Mays, L. E. (1990). Signal transformations required for the generation of saccadic eye movements. *Annual review of neuroscience*, *13*(1), 309-336.
- SR Research (2017). EyeLink 1000 Plus User Manual 1.0.12. Retrieved from <https://www.sr-research.com/filedata/fetch?id=49224>.
- Sprague, J. M. (1966). Interaction of Cortex and Superior Colliculus in Mediation of Visually Guided Behavior in the Cat. *Science*, *153*(3743), 1544-1547.
<https://doi.org/10.1126/science.153.3743.1544>
- Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 137-142.
- Stein, B. E. (1984). Development of the superior colliculus. *Annual Review of Neuroscience*, *7*, 95-125.
- Steinman, R. M. (1965). Effect of Target Size, Luminance, and Color on Monocular Fixation. *Journal of the Optical Society of America*, *55*(9), 1158-1164.

- Trevarth, C. B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung*, 31, 299–337.
- Tusa, R. J., Mustari, M. J., Burrows, A. F., & Fuchs, A. F. (2001). Gaze-Stabilizing Deficits and Latent Nystagmus in Monkeys With Brief, Early-Onset Visual Deprivation: Eye Movement Recordings. *Journal of Neurophysiology*, 86(2), 651–661. <https://doi.org/10.1152/jn.2001.86.2.651>
- Vidyasagar, T. R. (1978). Possible plasticity in the rat superior colliculus. *Nature*, 275(5676), 140.
- Wallace, M. T., McHaffie, J. G., & Stein, B. E. (1997). Visual Response Properties and Visuotopic Representation in the Newborn Monkey Superior Colliculus. *Journal of Neurophysiology*, 78(5), 2732–2741. <https://doi.org/10.1152/jn.1997.78.5.2732>
- Worfolk, R., & Abadi, R. V. (1991). Quick phase programming and saccadic re-orientation in congenital nystagmus. *Vision Research*, 31(10), 1819–1830.
- World Health Organization. (2004). ICD-10: international statistical classification of diseases and related health problems: tenth revision.
- Yee, R. D., Wong, E. K., Baloh, R. W., & Honrubia, V. (1976). A study of congenital nystagmus: Waveforms. *Neurology*, 26(April), 326–333.

APPENDIX

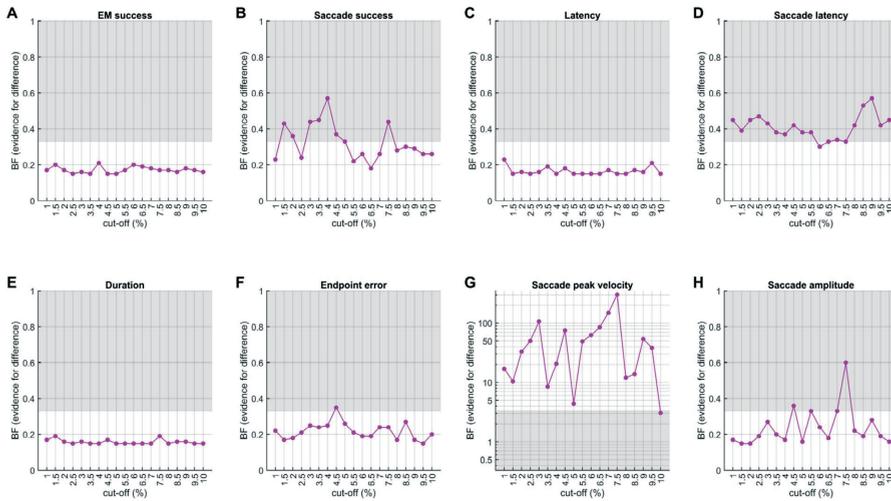


Figure A-1: Bayes factors (BF) indicating evidence for a difference in eye movement characteristics (see Fig. 3) between the CC and the NC group for 19 different cut-off values implemented in the classification algorithm (see *Data Analysis*). Grey shaded areas indicate BFs between 0.33 and 3, which represent inconclusive evidence.

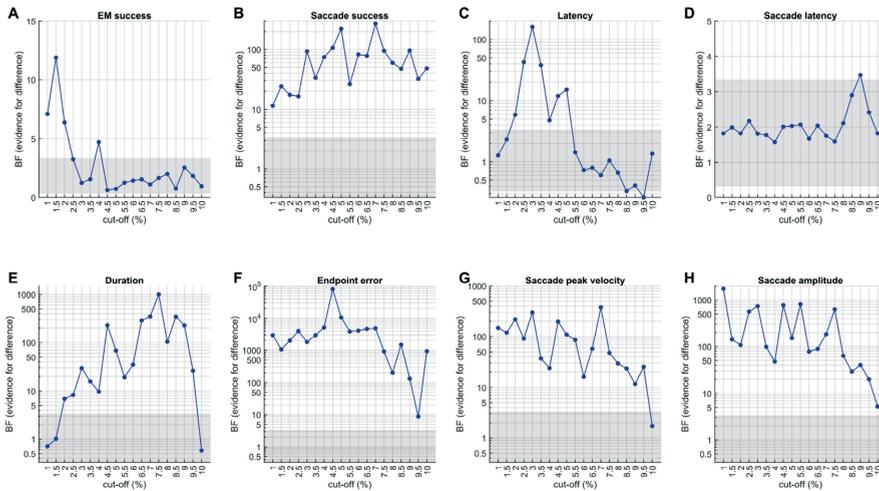


Figure A-2: Bayes factors indicating evidence for a difference in eye movement characteristics (see Fig. 3 in the main text) between the CC and DC group for 19 different cut-off values underlying the classification algorithm (see *Data Analysis*). Grey shaded areas indicate BFs between 0.33 and 3, which represent insufficient evidence.

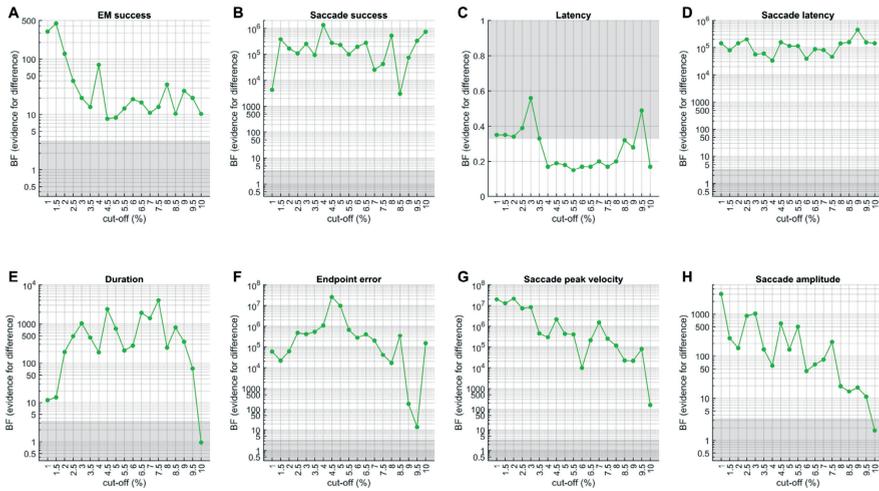


Figure A-3: Bayes factors indicating evidence for a difference in eye movement characteristics (see Fig. 3 in the main text) between the CC and SC groups for 19 different cut-off values underlying the classification algorithm (see *Data Analysis*). Grey shaded areas indicate BFs between 0.33 and 3, which represent insufficient evidence.

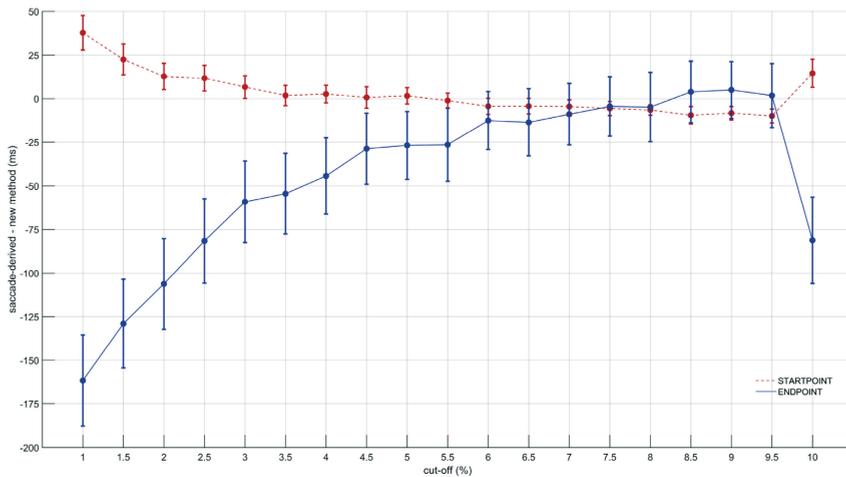


Figure A-4: Differences between start-and endpoints defined by the eye tracker based method (saccade-derived method) vs the newly suggested method in the sighted control group for 19 different cut-off values underlying the classification algorithm (see *Data Analysis*). Positive values indicate that the new method estimated start-and endpoints earlier than the saccade-derived method. Error bars indicate standard error of the mean. At a cut-off value of 7.5% had the standard and new method matched best.

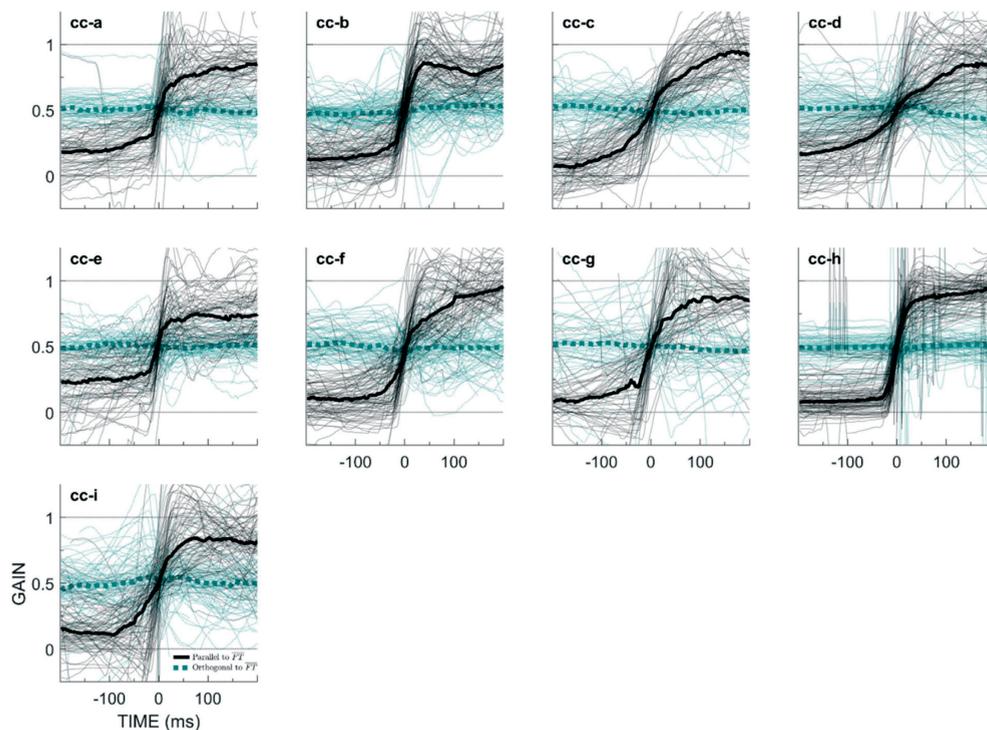


Figure A-5: Eye movement traces for each individual in the congenital cataract reversal group. Eye movement traces start 200 ms before and end 200 ms after a detected midpoint and indicate gaze position after transformation into two components: displacement parallel (grey solid lines) and orthogonal (cyan dotted lines) to the axis between the fixation target and saccade target. Traces were centred on the eye movement midpoint of each trial and were normalized to the distance between fixation target and saccade target. Horizontal dotted lines represent the normalized fixation target (0) and saccade target position (1). Individual plots show position traces for all trials with a detectable midpoint (thin lines) and the corresponding median trace of a participant (median position at each time point; thick lines). See also Fig. 2 in the main text.

2.4 Successful visually guided eye movements following sight restoration after congenital cataracts

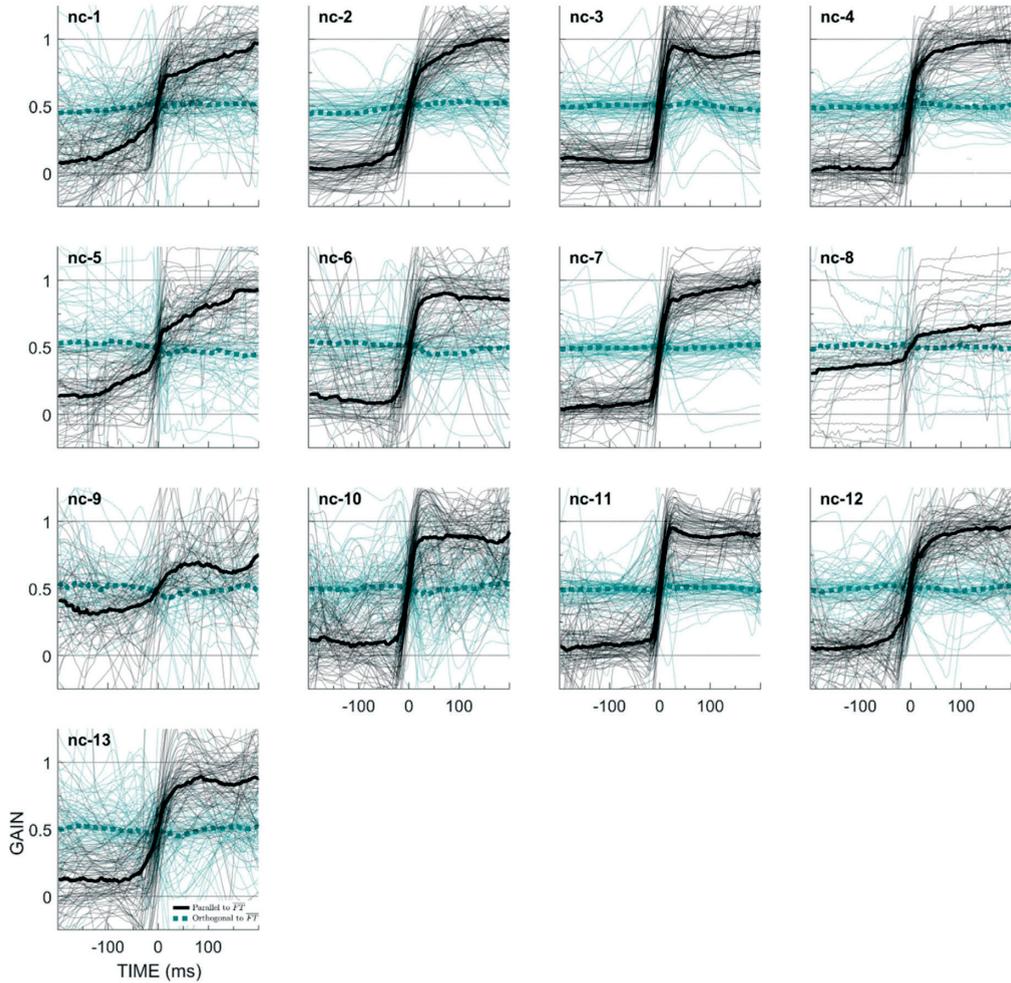


Figure A-6: Eye movement traces for each individual in the nystagmus group. See also Fig. A-5 and Fig. 2 in the main text.

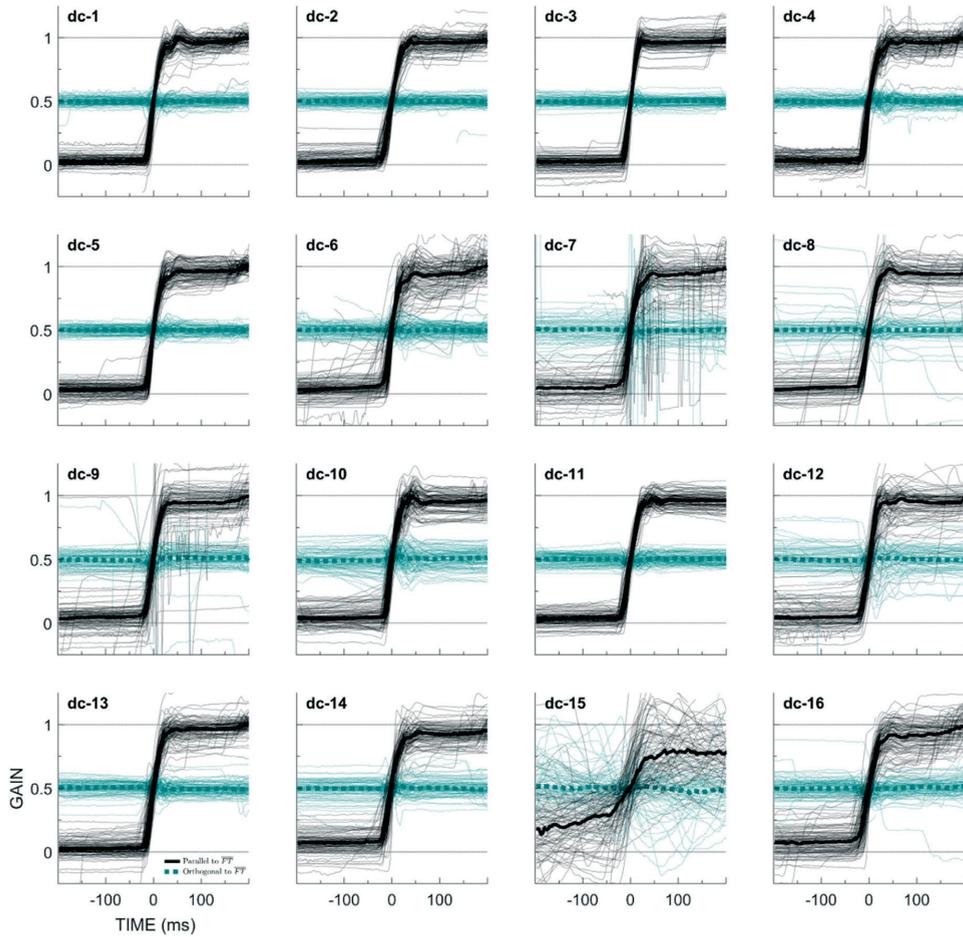


Figure A-7: Eye movement traces for each individual in the developmental cataract reversal group. See also Fig. A-5 and Fig. 2 in the main text.

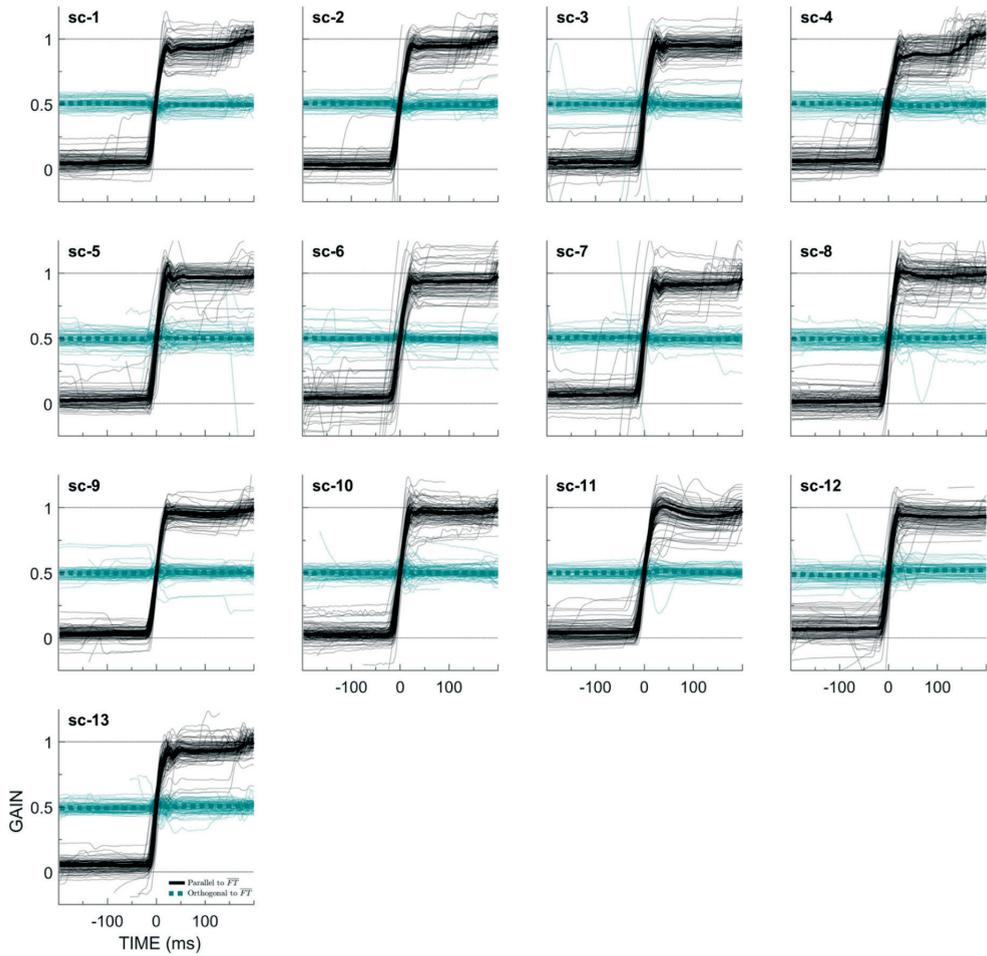


Figure A-8: Eye movement traces for each individual in the normally sighted group. See also Fig. A-5 and Fig. 2 in the main text.

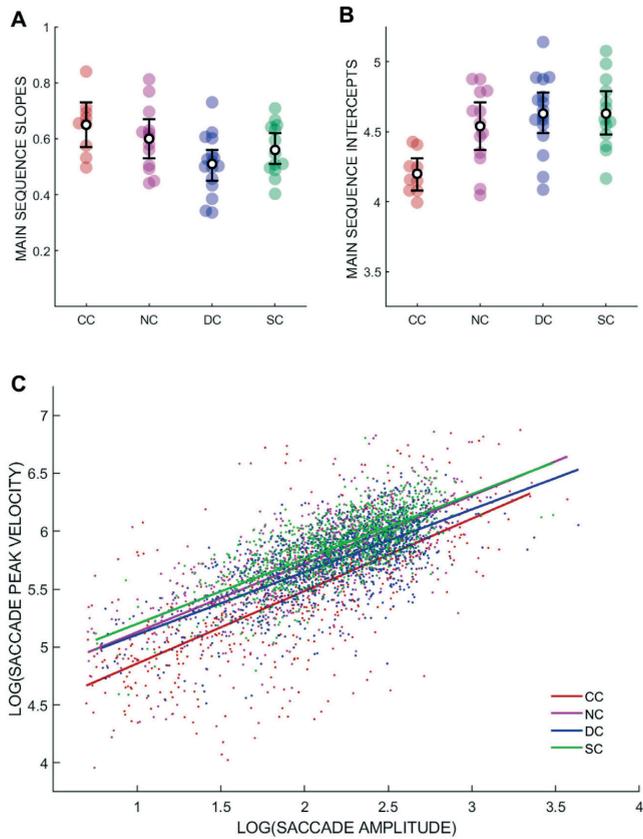


Figure A-9: Results of the saccade main sequence analysis. A) Slopes. B) Intercepts. Coloured, filled circles represent individual subject values; black, open circles represent estimated posterior distribution means per group, and bars represent 95% highest density intervals (HDI). C) Visual display of linear regression fits on all saccades per group. CC = congenital cataract group; NC = nystagmus control group; DC = developmental cataract group; SC = normally sighted control group.

2.4 Successful visually guided eye movements following sight restoration after congenital cataracts

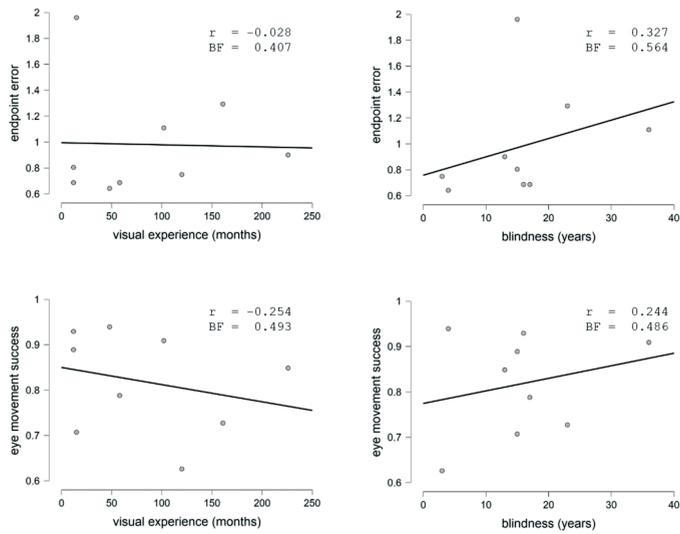


Figure A-10: Correlations between duration of visual experience, duration of blindness and endpoint error, eye movement success.

2.5 Serial dependence reflects integration of task-relevant information in working memory

ABSTRACT

Sensory input is inherently noisy while the world is comparatively stable and predictable. Optimal integration (or averaging) of multiple perceptual episodes may increase behavioural effectiveness by minimizing sensory measurement noise. For example, visual history effects indicate that behavioural reports are not just based on current observations but are also influenced by previous observations. In this study we demonstrate across four experiments (total $N = 332$) the crucial role of working memory in sequential response biases. We presented observers with a sequence of four oriented gratings and cued them to report one of these orientations. Reports were biased not only by distractors presented before the target, but also by distractors presented after target offset, indicating that reports are influenced by all relevant information currently held in working memory. We further show that behavioural reports are attracted towards distractors that are task-relevant and resemble the target on a relevant feature dimension and are repulsed by task-irrelevant and dissimilar distractors. Together, these findings suggest that memory representations are purposefully integrated according to the behavioural goals of the observer. We propose that these results reflect a mechanism that integrates information sampled from the same object at different timepoints in order to promote perceptual stability and behavioural effectiveness. Data and analysis scripts can be found at <https://osf.io/rmdqz/>.

Submitted as:

Zerr, P., Gayet, S., Van der Stigchel, S., (under review). Serial dependence reflects integration of task-relevant information in working memory

All authors designed the study concept. PZ programmed the experiment, collected experiment data, conducted the analyses, and wrote the manuscript. Critical revisions were provided by all co-authors.

INTRODUCTION

The macroscopic world is highly continuous. The features of our surroundings remain relatively stable within the time frame of seconds (e.g., van Bergen & Jehee, 2019, Figure 2) and objects tend to move in a predictable way. At the same time, sensory systems do not acquire perfect observations of the world in each instant. Instead, we typically rely on incomplete information from e.g., short gaze fixations and peripheral vision, to form an actionable image of the world. While driving a car, for example, we often estimate the speed and position of other cars from brief glances as we cannot afford to look in the mirror for long periods of time. Even without temporal constraints, perception is characterized by repeated, brief sampling behaviour as indicated by the average human fixation duration of only a few hundred milliseconds between saccades (Manor & Gordon, 2003). Yet, we are remarkably effective in our interaction with the world, suggesting the presence of adaptive mechanisms that account for noisy sensory input so that precise action output can be produced. For example, it is well established in multisensory perception research that information from different senses is combined in a statistically optimal way to form a more reliable estimate of a world state (Ernst & Banks, 2002; Alais & Burr, 2004). The same principle could account for noisy information sampling within a single sensory modality by combining multiple consecutive episodes of sensory input in a principled way to generate an increasingly reliable object representation. Averaging over multiple observations of same object necessarily reduces noise and leads to a more accurate estimate of a world state than any single perceptual estimate. In this study we investigate the proposition that behavioural reports of a world state rely on a weighted average (or functional integration) of memory representations created during consecutive episodes of sensory input.

Mechanisms in which past information influences current perception and action have been observed in phenomena such as priming (Tulving & Schacter, 1990), sensory adaption (Kohn, 2007), proactive interference (Makovski et al., 2008; Kiyonaga et al., 2017) and serial dependence (J. Fischer & Whitney, 2014; Burr & Cicchini, 2014). Serial dependence refers to a systematic bias, which can be observed in behavioural responses that are influenced by previously encountered information. For example, the reported orientation of a line may be consistently pulled towards the orientation reported in the previous trial. In recent years, serial dependence has been observed in a wide range of visual features such as orientation (J. Fischer &

Whitney, 2014; Liberman et al., 2016), colour (Barbosa & Compte, 2020), spatial position (Manassi et al., 2018), odour perception (Van der Burg et al., 2021), ensemble perception (Manassi et al., 2017), as well as high level features such as the emotional expression of faces (Kondo et al., 2012; Liberman et al., 2014), and attractiveness (Taubert et al., 2016; Xia et al., 2016). The process underlying serial dependence has been described as a predictive mechanism, in which previously encountered information aids current estimates (Cicchini et al., 2018; J. Fischer & Whitney, 2014; Fritsche et al., 2017, 2020). In other words, serial dependence may arise from an integration of previously and currently observed information in order to capitalize on the learned stability of the world and ameliorate the relative unreliability of individual perceptual events (i.e., the continuity field hypothesis; (J. Fischer & Whitney, 2014). Indeed, there are several reports of how previously observed information alters the perception of a new stimulus (Cicchini et al., 2017, 2021; Collins, 2020; Manassi et al., 2018). Some researchers have proposed a Bayesian and efficient coding model, in which a previous observation provides a prior, which is optimally combined with new incoming sensory information (Cicchini et al., 2018; Fritsche et al., 2017, 2020). In line with this view, it has been shown using generative modelling of fMRI data, that the influence of a previous stimulus on current stimulus report is strongest when the previous stimulus is represented with higher reliability than the current stimulus (van Bergen & Jehee, 2019). There is also evidence that human observers learn and apply knowledge of the statistics of the world in longer timeframes, spanning years, as well as very short timeframes, spanning just a few trials (Chetverikov et al., 2019; Dogge et al., 2019). Together, these results demonstrate that response biases such as serial dependence reflect behaviourally beneficial processes.

Investigations of serial dependence have focused on paradigms in which current perception, memory, and action are influenced by past information. This is typically done by examining the influence of a previous trial on the behavioural report in the current trial. In the present study we instead consider the hypothesis that behavioural reports of a given stimulus depend on a weighted average of consecutive episodes of sensory input that are stored in memory. To address this possibility, observers were presented with a sequence of four orientations and were cued to report one of them (Fig. 1). If stimulus representations are biased during memory maintenance (and not only when they are encoded), then any information that is concurrently maintained in memory should influence target report.

This leads to the specific prediction that target report is influenced not only by one or more stimuli presented *before* the target, but also by stimuli that are presented *after* the target has already disappeared (i.e., after encoding was completed). To preface our findings, stimuli presented after target offset also influenced target report. This implicates that behavioural responses to observed information are influenced by all relevant information that is concurrently stored in working memory. We demonstrate that response biases do not arise exclusively during perception or encoding, but can arise during memory maintenance or report, as a consequence of memory averaging.

We further investigated under which conditions response biases arise and show that this temporal memory averaging does not weigh all items equally. Items are weighted according to similarity, reliability, and task-relevance. Taken together, these findings show that consecutively presented stimuli are integrated in memory to serve behaviour. We discuss similarities and differences between the classic serial dependence bias and the temporal memory averaging observed in the present study.

Term	Definition
Orientation	Short for: orientation of an oriented sine grating with a Gaussian envelope, also known as Gabor patch.
Target	An orientation that was retro-actively cued for explicit report in a reproduction task.
Distractor	An orientation presented in the same trial as the target, which needed to be memorized but was not cued for report.
Response bias	A systematic deviation of behavioural responses relative to other observed stimuli. Here used to denote the bias under investigation.
Serial dependence	The response bias towards (attractive bias) or away (repulsive bias) from a previously observed stimulus. Also known as sequential response bias or visual history effect.
Target-Distractor difference; Δ	The orientation of a non-target item relative to the target orientation, with negative values indicating that the distractor was oriented counter-clockwise relative to the target, and positive values indicating clockwise relative orientation.

Table 1: Definition of terms used in this article.

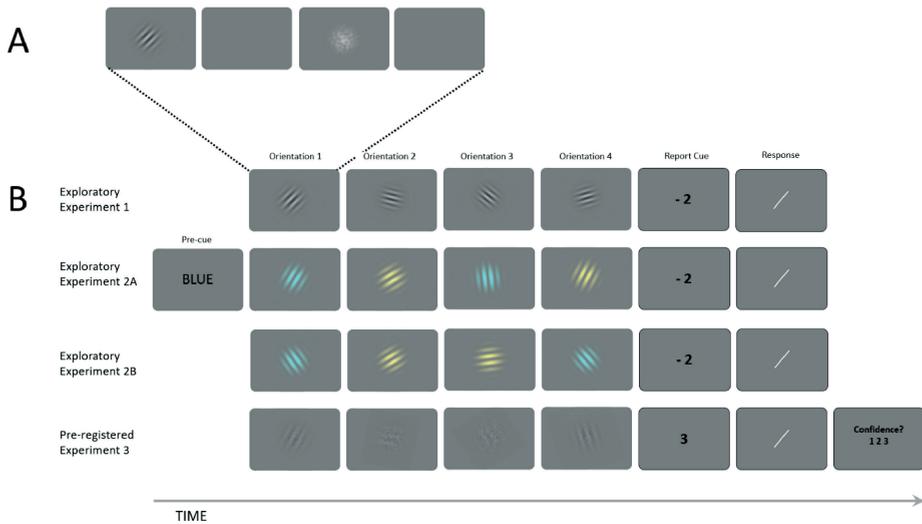


Figure 1: Schematic visualizations of experimental procedures. Sizes and colours are not to scale. A) The sequence of one orientation presentation, consisting of an orientation, a blank, a noise mask and another blank. B) Example trial sequences (excluding blanks and noise masks). In all experiments, observers were sequentially presented with four orientations and were cued to reproduce one of them, by means of an Arabic numeral. The reproduction of the cued orientation was enabled via a rotating white line. See Methods for a detailed description. In Experiment 1, standard orientation patches were used to present orientations. In Experiment 2A and B, orientations could be presented in one of two colours. In Experiment 2A, the colour of the upcoming target was pre-cued by the word blue or yellow, while in Experiment 2B no pre-cue was presented. In Experiment 3, orientations could be presented as high-noise or low-noise stimuli and observers were additionally asked to rate their confidence in the response via a button press.

RESULTS

We examined whether the reported orientation of a retroactively cued target was influenced by the preceding or succeeding items in a sequence of four items. Specifically, we measured the extent to which target report was biased towards or away from each of the three non-target items (i.e., distractors) in the sequence. Three hypotheses (outlined below) generated through modelling work on the data of an exploratory experiment (Experiment 1, $N = 43$) were pre-registered (<https://osf.io/ytfm9>), and then tested in an independent data set (Hypotheses 1, 2 & 3). Hypothesis 4 was theory driven and pre-registered, but not based on prior modelling work. All analyses reported in this section pertain to the data set of the pre-registered Experiment 3 ($N = 150$). Throughout the Results section (unless otherwise specified), p-values reflect the probability of a Type 1 error, estimated by comparing the observed data to 10,000 generated null-distributions (i.e., permutation tests; see Methods section for details). It should be noted that, across trials, the occurrence of so-called swap errors (instances in which a distractor orientation was reported instead of the target) can resemble an attractive bias, as in both cases target report will appear to be pulled toward distractor orientations. Here, we identified and removed trials with swap errors using the Bays lab working memory modelling toolbox (Bays et al., 2009).

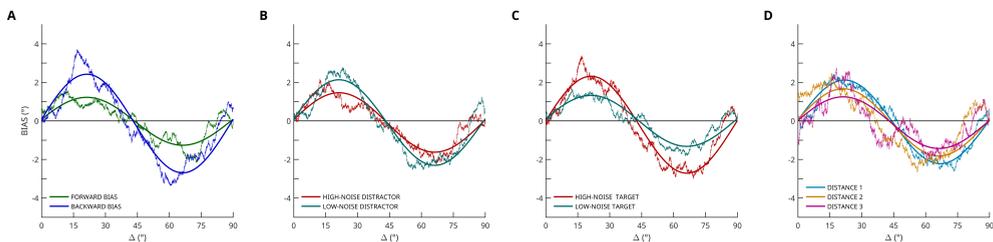


Figure 2: Results of the pre-registered Experiment 3. (A) Forward (distractor shown before target) and backward (distractor shown after target) response bias obtained from sorting response errors to the targets, as a function of the orientation difference between the target and a given distractor. Ragged lines represent moving averages over response errors. Smooth lines represent a fitted sine function. (B) Response bias for high-noise and low-noise distractors. (C) Response bias for high-noise and low-noise targets. (D) Response bias for distractors one, two or three items removed from the target in the sequence.

Working memory representations are biased by both preceding and succeeding items

If response biases such as serial dependence can arise during memory maintenance, then information encountered *after* memorization (and after target offset) should influence target report as well (Hypothesis 1). We found that orientation reports were not only biased by items that preceded the target (forward bias, consistent with the classic serial dependence effect; $p < .0001$; Fig. 2A), but also by items that succeeded the target (backward bias; $p < .0001$; Fig. 2A). This clearly demonstrates that response biases in this paradigm arise during working memory maintenance, because the backward bias can only emerge after the target stimulus was removed from the screen. This response bias emerged as both an attractive bias for small target-distractor differences, and a repulsive bias for large target-distractor differences. This was observed for both forward and backward biases (Fig. 2A), but surprisingly, the magnitude of the backward response bias was larger than that of the forward bias ($p < .0001$; difference between conditions; exploratory analysis).

Distractors in both close and far temporal proximity to the target induce a response bias

If the bias arises in memory, not just stimuli immediately preceding and succeeding the target item, but stimuli in the sequence that are one or two items removed from the target, should influence target report as well (Hypothesis 2). The data show that both close and far distractors biased target report ($p < .0001$ temporal distance 1; $p < .0001$; temporal distances 2&3), providing strong evidence that not just distractor orientations observed immediately preceding or succeeding the target item induce a response bias, but also target items further away in time.

We also expected distractors closer in time to the target to induce a stronger bias and distractors further from the target to induce a weaker bias (Hypothesis 3). However, the data revealed only a trend in the predicted direction. A permutation test with shuffled condition labels revealed no significant difference in bias magnitude between neighbouring and more distant distractors ($p = .0583$). Fig.

2D displays the response bias for each of the three temporal distances separately.

Stronger bias for less reliable targets and more reliable distractors

One functional benefit of generating an averaged representation from multiple observations could be noise reduction: averaging multiple noisy orientation estimates should lead to a less noisy orientation representation. Less reliable representations should be weighted less, and more reliable representations should be weighted more. Specifically, low-noise distractors should influence target report less than high-noise distractors (Hypothesis 4A), and low-noise targets should be more strongly influenced by distractors than high-noise targets (Hypothesis 4B).

The data revealed a difference in bias magnitude between low-noise and high-noise distractors ($p = .0245$; Fig. 2B), and between low-noise and high-noise targets ($p = .0003$; Fig. 2C). This finding suggests that memory averaging exerts a stronger influence on memory reports, when the to-be-reported item contains more noise; that is, when averaging would be most beneficial to behaviour. Conversely, when distractors were more reliable (less noisy), they influenced target report more strongly than when they were less reliable (more noisy). As the probability threshold for identifying swap errors had not been pre-registered, permutation test p-values are plotted as a function of swap error threshold in Appendix Fig. 5. The results are reported for a probability threshold of 0.4 but hold for a range of reasonable values.

Exploratory Analyses

Unless otherwise specified, the analyses reported here are conducted on the combined data sets of Exploratory Experiment 1, Experiment 2B and Pre-registered Experiment 3 (total $N = 248$).

All items in working memory are influenced by every other item

If behavioural responses are the result of a functionally weighted average representation obtained during separate episodes of sensory input before making a behavioural decision, then the

reproduction of a memorized item should be influenced by all similar items in working memory. That is, most (if not all) items in the sequence should influence the representation of every other item in the sequence, provided they are stored in memory. Indeed, all combinations of distractor and target produced a significant response bias (all $p < .0001$; Fig. 3), causing attraction for similar orientations ($\Delta < 45^\circ$) and repulsion for dissimilar orientations ($\Delta > 45^\circ$).

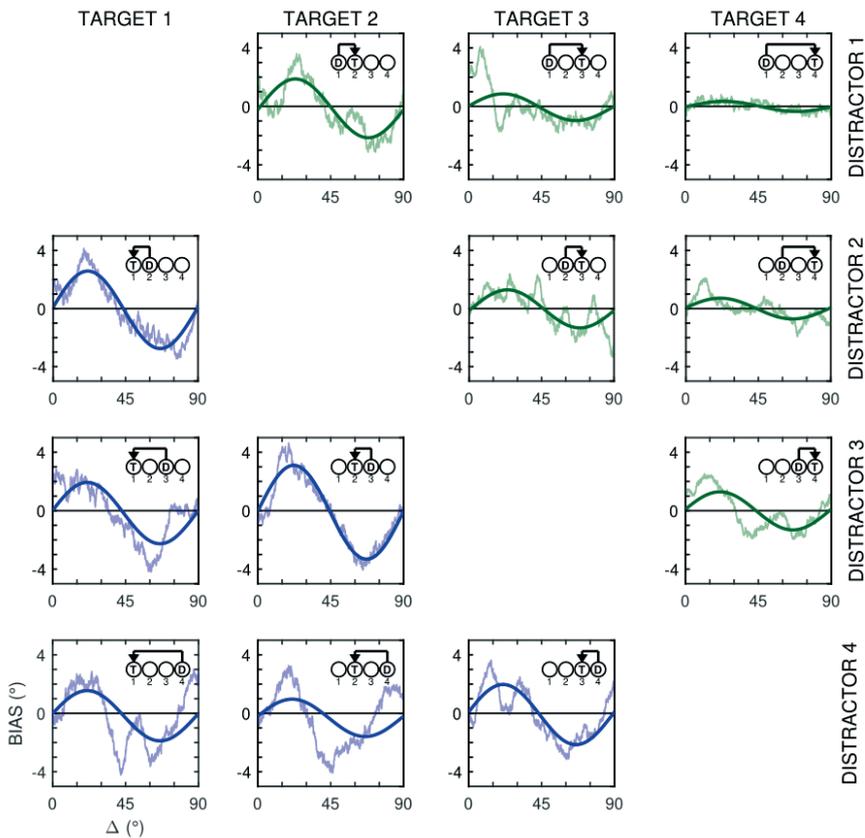


Figure 3: Exploratory results on the combined datasets. Influence of distractor orientation on target orientation report, for each target-distractor combination. The ragged lines represent a moving average of response errors (i.e., response bias; on the y-axis) sorted by the orientation difference between target and distractor (i.e., distractor difference; on the x-axis). The insets in the top right of each panel indicate the analyzed target-distractor combination. The smooth lines depict a fitted sine function.

Opposite effects of task-relevant and task-irrelevant distractors

A functionally weighted average should include only those items that are task-relevant. If the response bias reported here reflects a functionally beneficial principle, rather than an artefact of the system, we would expect that only behaviourally task-relevant items in memory influence target report. To investigate how task-relevance impacted the response bias, we conducted Experiments 2A and 2B, in which each orientation in the sequence was presented in one of two colours.

In Experiment 2A observers were pre-cued with one of the two colours and instructed that only targets of the pre-cued colour would be tested. This manipulation produced a dissimilarity between memorized items on a task-relevant feature dimension. We found a response bias consistent with the data pattern observed in Experiments 1 and 3 when the distractor was the same colour as the target (Fig. 4A). Target reports were attracted toward similar distractors ($\Delta < 45^\circ$), whereas target reports were repulsed by dissimilar distractor orientations ($\Delta > 45^\circ$). In contrast, when target and distractor were of a different colour, we observed a reversal of this effect: similar distractor orientations caused a repulsion of the target report ($p < .0001$). The data revealed a significant difference between these two conditions ($p < .0001$). Beyond a reversal of the data pattern as modelled by a sine function, visual inspection of the model-free moving average in Fig. 4A suggests that unattended distractors (i.e., distractors of a different colour) resulted in a response bias pattern exactly opposite that of attended distractors, at least for target-distractor differences smaller than 60° . The repulsion effect for dissimilar colours suggests that items that were known to the observer to be task-irrelevant as they were not going to be tested and did not need to be memorized, were nonetheless stored in memory, and influenced responses in the opposite direction from the influence of task-relevant working memory items.

The result of Experiment 2A might have been caused simply by the dissimilarity between differently coloured items. To rule out this explanation we conducted Experiment 2B, in which observers were presented with the same, coloured orientations as in Experiment 2A, but the target colour was not pre-cued. In contrast to Experiment 2A, no difference in response bias was found between differently and same coloured target-distractor pairs ($p = .9347$). Instead, the response bias as observed in Experiments 1 and 3 was present when targets and

distractors were of same colour ($p < .0001$) and when they were of different colours ($p < .0001$). This result indicates that dissimilarities between distractor and target on a task-relevant, but not task-irrelevant feature dimension reduce or even reverse the response bias.

Self-reported confidence and serial position predicts response bias magnitude

After each trial in Experiment 3 observers rated their confidence in the orientation report that they just provided. We expected that orientation reports on low-confidence trials (“Not quite sure.”) would be influenced more strongly by distractors than on high-confidence trials (“I am reasonably sure.”). Self-reported guesses (“Pure guess, I have no idea.”) were excluded from this analysis. This is indeed what the data show ($p = .0339$; comparison between conditions; Fig. 4C), indicating that observers had knowledge of the reliability of their memory representations and adding support to Hypotheses 4A, which states that less reliable targets are influenced more strongly by distractors.

Targets earlier in the sequence are influenced more strongly

Target reports for items earlier in the sequence are characterized by a decreased memory recall precision (see also Appendix Fig. 6, showing error distributions for responses to each of the four targets separately). This is to be expected as representations decay over time and become more variable as additional information is encoded. Thus, targets earlier in the sequence should be more strongly influenced by distractors. This is exactly what we found. We computed to what extent response errors on the target orientation reports were biased toward the combined influence (i.e., circular mean) of the three distractor orientations. Comparing the response bias magnitude in the first and last target in the sequence (using datasets from Experiment 1, 2B and 3) reveals a significant difference in the predicted direction: the first target in the sequence was more strongly influenced than the last target in the sequence ($p = 0.0281$; difference between conditions; Fig. 4D). This finding adds support for Hypotheses 4A, indicating that the less reliable, earlier targets were influenced more strongly than more reliable, recent targets.

Previous responses bias current responses

In the previous sections we focused on within-trial effects, that is, the influence of representations currently held in working memory on target report. Conveniently, the present paradigm allows to concurrently examine the influence of the previous trial on current target report. Previous target response strongly biased current responses ($p < .0001$). This represents a replication of the classic serial dependence effect (Fig. 4E). The tuning curve of the response bias from previous trial responses (between-trial effect) bears remarkable similarity to that of the within-trial effects reported in the rest of this study (Fig. 4F). In addition, the veridical (i.e., displayed) orientation of the target in the previous trial also biased current responses ($p < .0001$), but to a lesser extent than the previous responses ($p < .0001$; difference between conditions).

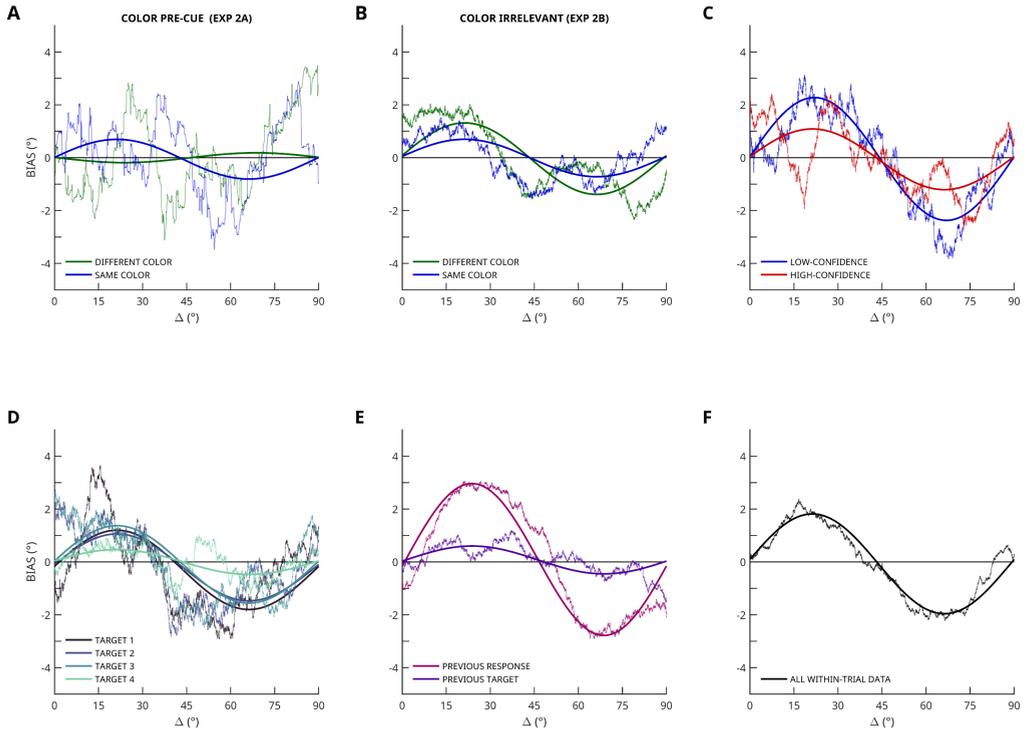


Figure 4: Exploratory results. (A) Response bias for trials in which target and distractor had the same or different colours and the colour of the test item was pre-cued. (B) Response bias when target and distractor had the same or different colours and colour was not task-relevant. (C) Response bias for trials with self-reported low or high confidence. (D) The combined influence of the three distractors (i.e., circular mean) on each of the four targets in the sequence. (E) The influence of the previous target and the previous response. (F) The response bias for all datapoints from Experiment 1, 2B and 3 combined. Ragged lines represent moving averages. Solid lines represent a fitted sine function.

DISCUSSION

In a relatively stable and continuous world, any system making inferences about a world state that is faced with the inherent unreliability of sensory input benefits from integrating information from multiple sources. The average representation of several relatively unreliable observations necessarily produces a more accurate estimate of a world state. Most prominently, this principle has been demonstrated for multisensory integration (Ernst & Banks, 2002; Alais & Burr, 2004), and emerges in the present study as a functional integration of multiple memory representations within the same modality. We investigated how the reproduction of a memorized target, cued from a sequence of four orientations, was systematically biased by the three non-targets (distractors). The results show that target report was consistently influenced by orientations that were shown *before* the target in the sequence as well as orientations shown *after* the target was no longer visible. Responses were influenced more strongly when the target representation was less reliable or the distractor representation more reliable. Our findings reveal that not just the target, but any similar information that is currently held in working memory influences target report. Our data further revealed that responses are systematically attracted toward the orientation of all stimuli in the sequence, provided they are sufficiently similar on a task-relevant feature dimension. We propose that response biases not only reflect the influence of single distractor representations but arise from a functionally weighted average of target and distractors observed across multiple episodes of sensory input.

Memory representations are biased by previously and subsequently memorized stimuli

In addition to a “forward bias” (i.e., similar to classic serial dependence), we also found a “backward bias”. Distractors biased target response even when stimuli were presented after perceptual and encoding stages of the target were completed (Hypothesis 1). It has been a subject of recent debate whether response biases such as serial dependence arise during perception, that is, whether it is the perception of a new stimulus that is biased (Cicchini et al., 2017, 2021; Collins, 2020; Feigin, 2021; Fornaciai & Park, 2018; Manassi et al., 2018; St. John-Saaltink et al., 2016), or whether the bias arises during post-perceptual stages, such as memory maintenance (Bae & Luck, 2017; Barbosa & Compte, 2020; Bliss et al., 2017; Ceylan et al., 2021;

Fritsche et al., 2017). It is reasonable to assume that integration mechanisms such as serial dependence are implemented at different processing stages (Ceylan et al., 2021; Pascucci et al., 2019). While the present results cannot rule out that biased representations also arise during stimulus encoding, they do establish that representations can be biased after stimulus encoding. A response bias among memorized items has previously been shown for random dot motion stimuli (Fornaciai & Park, 2020). The authors presented sequences of three dot motion displays and then cued observers to report the motion direction in one of them. The data revealed that in some cases subsequently presented motion directions attracted target report of a previously presented dot motion display. Here, we extend this finding by investigating under which conditions this response bias arises, and which factors modulate its magnitude.

Functional memory integration is weighted by similarity, reliability, and task-relevance

One crucial benefit of integrating memory representations could be noise reduction. Behavioural effectiveness crucially depends on a functionally optimal representation of world states. Any estimate based on a functional integration of several observations necessarily is more reliable than a single observation. Recent work has shown that sequentially presented stimuli are simultaneously represented in brain activity even without explicit memorization instructions (King & Wyart, 2021). Such concurrent storage potentially allows the visual system to combine multiple observations to generate an averaged representation that is more accurate than any single observation alone. However, it would also be expected that not all observations are weighted equally. For example, a behaviourally beneficial memory integration mechanism would be expected to integrate only those representations that are likely to have been sampled from the same object. Indeed, it became evident in the present data that target report is consistently attracted by *similar* orientations in memory ($\Delta < 45^\circ$) and repulsed by dissimilar orientations ($\Delta > 45^\circ$). Similar effects of target-distractor similarity were previously reported in the context of visual search (Rafiei et al., 2021).

Another expected attribute of behaviourally beneficial memory integration would be that more *reliable* information is weighted more strongly than unreliable information. Our data show that noisy, less reliable targets were more influenced by distractor items (Hypothesis

4A) as compared to less noisy targets. Conversely, less noisy, more reliable distractors exhibited a stronger influence on target responses than more noisy, less reliable distractors (Hypothesis 4B). These findings demonstrate that when individual observations are unreliable, the visual system may rely more heavily on information from other observations of the same object. In line with our current findings, Van Bergen & Jehee (2019) showed that the influence of unreliable on more reliable representations was smaller than the influence of reliable on unreliable representations. Furthermore, it has recently been demonstrated that uncertainty is maintained and used in working memory and reflects the precision of memory item representations (Yoo, Ascerbi & Ma, 2021). Our data extend these findings in showing that the magnitude of the bias is influenced by the reliability of the target, as well as the reliability of the distractor.

When information from multiple observations is combined, it would be beneficial to weigh these observations according to their *relevance* to the behavioural goal. It has previously been shown that when conjunct objects had to be remembered, task-relevant features induced a larger serial dependence bias than task-irrelevant features (C. Fischer et al., 2020). In the present study, when presented with orientations randomly coloured in one of two colours, observers' responses were biased toward all other distractors, irrespective of their colour. However, when observers were additionally presented with an informative pre-cue before the item sequence, indicating the colour of the target, non-cued colours displayed a response bias pattern exactly opposite to that of cued items (Fig. 4A). The presence of a significant, reversed response bias indicates that un-cued distractors were encoded, and their representations actively distinguished from the target.

Functional memory integration as mechanism underlying serial dependence

When a previous observation influences a current observation, it must have been stored in some form. Thus, any response bias is intimately linked to memory. Serial dependence is thought to reflect a very similar principle as memory averaging: by taking into account previously encountered information, presently encountered information can be more reliably estimated. The findings of the present study suggest that serial dependence might be part of a more general process of memory averaging over time (or temporal

smoothing). In contrast to typical serial dependence paradigms, in our study participants were explicitly instructed to memorize all stimuli in a sequence. In typical serial dependence paradigms, participants are only instructed to report the last seen stimulus, such that there is no explicit requirement to memorize the preceding stimulus. Could it be that distinct mechanisms underlie the classic serial dependence effect, and the memory-related response biases observed in the current study? For serial dependence to occur, information from the previous stimulus must have lingered in memory in some form despite no instruction or benefit of memorizing these stimuli beyond the trial. Indeed, our exploratory analysis showed that the response to the current target stimulus was also biased toward the response in the previous sequence, which participants no longer needed to keep in memory. The tuning function of this effect was strikingly similar to that of the observed within-trial biases (Fig. 4C), demonstrating that an explicit memorization instruction is not required to induce a response bias. A potential purpose of storing multiple observations in memory (e.g., King, & Wyart, 2021) could be functional integration of working memory. It is therefore very well possible that the mechanism underlying classic serial dependence also relies on the simultaneous memorization of target and inducer stimuli. Serial dependence might be part of a more general process of memory averaging over time to facilitate more efficient perception and action.

In summary, our findings show that response biases observed in the present study – as well as those previously observed in the literature – likely reflect different expressions of a generic mechanism of memory integration, in which individual representations are weighted according to their similarity on a task-relevant feature dimension (i.e., orientation, Experiments 1 to 3 and colour, Experiment 2A) and their reliability. This process may facilitate noise reduction in the face of sensory uncertainty and support the stability of perception and behaviour. The visual system effectively makes use of the statistical regularities and continuity present in the real world, on both larger and shorter timescales. As the present results show, working memory appears to play a key role in the implementation of these principles.

METHODS

Results reported here are based on four experiments. Modelling work based on Experiment 1 formed the basis of the pre-registered Hypotheses 1, 2 and 3, while pre-registered Hypothesis 4 was theory driven rather than based on prior modelling work. Experiment 3 was designed to test the four pre-registered hypotheses (Pre-registered Results section). In addition, Experiments 2A and 2B were conducted to address further questions of interest that were not pre-registered (Exploratory Results section). The pre-registration protocol can be found at <https://osf.io/ytfm9>.

Participants

Observers were recruited via the online research platform Prolific (www.prolific.co) and were compensated for their time with 8,- euro/hour. Inclusion criteria were set to age 23 - 50 years, no literary difficulties, completed education level of at least BA/BSc, no psychiatric medication use, normal or corrected to normal vision, fluency in English, no daily impact of mental illness, less than 10 units of alcohol consumed per week, at least 95% approval rating on Prolific and 10 minimum previous experiment participations.

Observers providing data with flat error distributions (indicating random responses) were excluded from analysis (Exp1: 19, Exp2A: 2, Exp2A: 17, Exp3: 10), as well as observers who failed an attention check (a simple question that could only be answered correctly by viewing the entire instruction video; Exp. 2B: 6, Exp. 3: 11). In addition, and according to the pre-registered protocol, 3 observers with extremely similar error distribution histograms between the first and the last target were excluded from Experiment 3 due to suspected cheating as it is highly unlikely that observers were able to perform equally well for the first (most difficult) and last (easiest) target. After exclusion, 43 observers participated in Experiment 1, 84 observers participated in Experiment 2A, 55 observers participated in Experiment 2B and 150 observers participated in Experiment 3.

Stimuli and Materials

All stimuli were displayed at the centre of the screen in sequence. The exact size and colour of stimuli varied according to the screen of the participant, which was required to be at least 719 px in height. Similarly, timing inaccuracies associated with online experiments likely resulted in some variance in presentation times (in the order of up to tens of ms). It is highly unlikely that the variance in stimulus presentation would have influenced the pattern of results reported here. There was always only one stimulus on the screen at a time and all manipulations were within-subject (except for the contrast between Experiments 2A and 2B). While such variation might have increased measurement noise, this would only result in more conservative test outcomes and any variation in presentation parameters would make the experiment more representative of the natural world. The screen background was set to 50% gray (RGB 127 127 127). Memory stimuli consisted of oriented line gratings (orientations), with a Gaussian envelope with a standard deviation of 80 pixels and a spatial frequency of 75 pixels per cycle. The mask, which was displayed after each orientation (a orientation patch, see Table 1), consisted of procedurally generated noise with a spatial frequency distribution matching that of the orientation presentations, but which contained no orientation information. In Experiment 2A and 2B, each orientation could be displayed as randomly drawn from one of two colours, cyan (RGB: 0 255 255) and yellow (RGB: 255 255 0). In Experiment 2A, an informative pre-cue was presented before the orientations, which consisted of the word “BLUE” or “YELLOW” at the centre of the screen. In Experiment 3, the noise manipulation was induced through a linear combination of the orientation stimuli and smoothed white noise. Each orientation could be displayed as randomly drawn from one of two noise levels: high-noise, with a stimulus contrast of 0.05 and a noise contrast of 0.4, or low-noise, with a stimulus contrast of 0.2 and a noise contrast of 0.3.

Procedure

Each experiment began with an instruction video, explaining the experiment to the observer, who were in addition instructed to find a quiet, dark, and clean space to conduct the experiment. For Experiments 2 and 3, at the end of the video a hint was given that the subsequent questionnaire item “What is 25+13?” should be answered with “19”. This was done to test whether an observer actually watched

the instruction video. Subjects who responded anything other than “19” (including the correct outcome of the equation, “38”) were excluded from analysis. This was followed by a practice round of 11 trials. After each practice trial, participants were given feedback on the magnitude of the response error in that trial. After the practice round, the main experiment began. After every 10 trials, feedback was given indicating the average response error over these last 10 trials. After 20 and 40 minutes experiment time (Exp. 1 and 2) or after 25 minutes (Exp. 3), observers were invited to take a short break. Experiments 1 and 2 terminated after 60 minutes total time, and Experiment 3 terminated after 45 minutes.

Experiment 1

Each trial began with a blank screen for 500 ms, followed by four repetitions of the sequence: orientation for 300 ms, blank for 200 ms, noise mask for 100 ms, blank for 750 ms and a cue for 200 ms. The cue indicated which of the four orientation targets the observer should report: -4 for the first orientation, -3 for the second, -2 for the third and -1 for the most recent orientation in the sequence. Following the cue was a blank for 750 ms, after which a randomly oriented line was presented, which could be rotated using the mouse to indicate which orientation was remembered. A mouse click confirmed the response and the next trial began. See Figure 1B for a visual depiction of a trial sequence in each experiment.

Experiment 2 A & B

Each trial began with a blank screen for 500 ms, followed by a pre-cue in Experiment 2A, consisting of the written-out words “BLUE” or “YELLOW” in black, which was omitted in Experiment 2B, followed by four repetitions of the sequence: orientation for 300 ms, blank for 200 ms, noise mask for 200 ms, blank for 750 ms. Each orientation and corresponding mask were colourized in one of two colours: cyan or yellow. This was followed by a cue for 200 ms. The cue indicated which of the four orientations the observer should report: -4 for the first orientation, -3 for the second, -2 for the third and -1 for the most recent orientation in the sequence. Following the cue was a blank for 750 ms, after which a randomly oriented line was presented, which could be rotated using the mouse to indicate which orientation was remembered. A mouse click confirmed the response and the next trial began.

Experiment 3

Each trial began with a blank screen for 500 ms, followed by four repetitions of the sequence: orientation for 200 ms on average (randomly drawn from a range between 140 and 260 ms, in steps of 8 ms), blank for 100 ms, noise mask for 50 ms, blank for 500 ms and a cue for 200 ms. The cue indicated which of the four orientation targets the observer should report: 1 for the first orientation, 2 for the second, 3 for the third and 4 for the most recent orientation in the sequence. Note that this is a reverse cue coding from Experiments 1 and 2, in order to make sure effects were not influenced by order of recall, i.e., whether participants went through their memory representations forward or backward. Following the cue was a blank for 750 ms, after which a randomly oriented line was presented, which could be rotated using the mouse to indicate which orientation was remembered. A mouse click confirmed the response. After responding observers were asked to rate their confidence in the accuracy of their response by pressing 1, 2 or 3 on the keyboard, corresponding to “Pure guess, I have no idea.”, “Not quite sure.” And “I am reasonably sure.”, respectively. These hints were displayed on screen, next to the corresponding number. After observers gave their confidence response the next trial began.

Data Analysis

All data and code as well as supplementary materials can be found at <https://osf.io/rmdqz/>.

Each trial contained three “data points”: one for each target-distractor combination. That is, there were three possible items in a trial that could bias report of the fourth item, the target. Each data point was represented by the angular difference between target orientation and reported orientation (response error), as well as the angular difference between distractor and target. In order to estimate the presence and magnitude of a response bias, response errors were sorted by target-distractor difference. Since no meaningful difference between clockwise and counter-clockwise orientations could be expected, these data were combined in such a way, that target-distractor differences are expressed in positive values (e.g. a difference of -80° becomes 80°) and represent the absolute angular difference Δ . Response errors corresponding to such sign-flipped target-distractor differences were also sign-flipped, such that an attractive bias is always expressed in positive, and a repulsive bias always expressed in

negative values of the response errors. A sine function was then fitted to this data through least-squares optimization (Equation 1; e.g., Fig. 2).

$$\text{Equation 1} \quad f(x) = \alpha \sin\left(x \frac{2\pi}{90} + \beta \frac{2\pi}{90}\right) + \gamma,$$

where x is response error in degrees, α is sine amplitude, β is the phase offset and γ is the vertical offset. The period of the sine was fixed to 90° . To ensure convergence, bounds were set to $\alpha = [-20 \ 20]$, $\beta = [-10 \ 10]$; $\gamma = [-3 \ 3]$. Starting values were set to $\alpha = 2$, $\beta = 0$, $\gamma = 0$.

Statistical analysis was performed through permutation tests. p -values were obtained by comparing the amplitude of the best fit on the observed data to the amplitudes of the best fits on 10,000 random permutations of the dataset. These permutations were produced by shuffling target-distractor differences. p -values indicate the proportion of shuffled data sets in which the amplitude of the best fit was equal to or larger than that of the ordered (i.e., observed) data set. An effect was considered to be significant at the $\alpha = 0.05$ level. Differences between conditions were obtained through similar permutation tests, but with shuffled condition labels instead of shuffled target-distractor differences (e.g., Ceylan et al., 2021; Fischer & Whitney, 2014).

Deviations from the pre-registered protocol

There were two deviations from the pre-registered protocol. First, trials with swap errors were removed from the analysis according to the modelling toolkit made available by the Bays lab (Bays et al., 2009). Swap errors (trials in which observers reported a different memory item than the cued target) would have strongly confounded the results, as a swap error data point would correspond to a 100% attractive bias. The inclusion of even a few swap errors would yield a pattern of response errors that is indistinguishable from a small but consistent attractive bias across all target-distractor differences. In our analysis, ~1000 of the ~15000 trials in Pre-registered Experiment 3 were classified as swap errors, that is, all trials that were classified with a swap error probability of 0.4 or higher. Following this exclusion, the data was revealed to not only contain an attractive bias, but both attractive and repulsive biases. Estimating

response bias magnitude according to the pre-registered protocol (i.e., the mean of the sorted response errors; see Data Analysis section above for details) would have produced a bias close to zero, as attractive and repulsive biases would have cancelled out. Instead, we fitted sine functions to the sign-flipped errors (see Data Analysis section above for details) and estimated bias magnitude as the amplitude of the sine function.

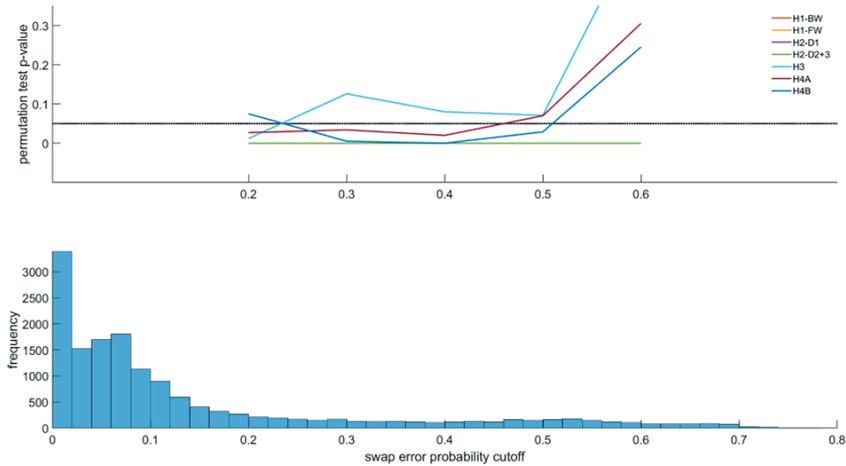
REFERENCES

- Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. <https://doi.org/10.1016/j.cub.2004.01.029>
- Bae, G.-Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception, & Psychophysics*, *79*(8), 2376–2395. <https://doi.org/10.3758/s13414-017-1404-8>
- Barbosa, J., & Compte, A. (2020). Build-up of serial dependence in color working memory. *Scientific Reports*, *10*(1), 10959. <https://doi.org/10.1038/s41598-020-67861-2>
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. <https://doi.org/10.1167/9.10.7>
- Bliss, D. P., Sun, J. J., & D’Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, *7*(1), 14739. <https://doi.org/10.1038/s41598-017-15199-7>
- Burr, D., & Cicchini, G. M. (2014). Vision: Efficient Adaptive Coding. *Current Biology*, *24*(22), R1096–R1098. <https://doi.org/10.1016/j.cub.2014.10.002>
- Ceylan, G., Herzog, M. H., & Pascucci, D. (2021). Serial dependence does not originate from low-level visual processing. *Cognition*, *212*, 104709. <https://doi.org/10.1016/j.cognition.2021.104709>
- Chetverikov, A., Hansmann-Roth, S., Tanrikulu, Ö. D., & Kristjánsson, Á. (2019). Feature Distribution Learning (FDL): A New Method for Studying Visual Ensembles Perception with Priming of Attention Shifts. In S. Pollmann (Ed.), *Spatial Learning and Attention Guidance* (Vol. 151, pp. 37–57). Springer US. https://doi.org/10.1007/7657_2019_20
- Cicchini, G. M., Benedetto, A., & Burr, D. C. (2021). Perceptual history propagates down to early levels of sensory analysis. *Current Biology*, *31*(6), 1245–1250.e2. <https://doi.org/10.1016/j.cub.2020.12.004>
- Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, *17*(14), 6. <https://doi.org/10.1167/17.14.6>
- Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1890), 20181722. <https://doi.org/10.1098/rspb.2018.1722>

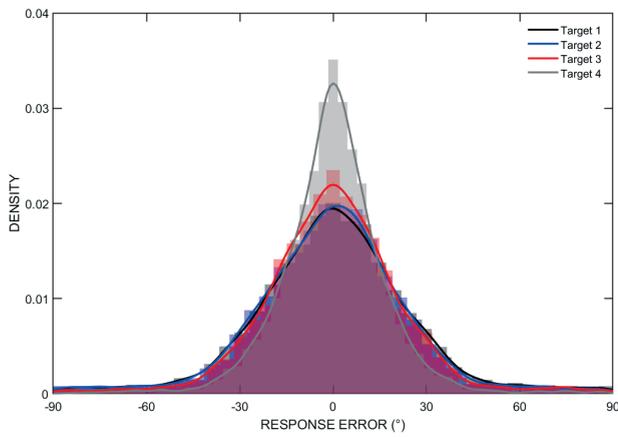
- Collins, T. (2020). *Serial dependence alters perceived object appearance* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/3gehk>
- Dogge, M., Custers, R., Gayet, S., Hoijtink, H., & Aarts, H. (2019). Perception of action-outcomes is shaped by life-long and contextual expectations. *Scientific Reports*, *9*(1), 5225. <https://doi.org/10.1038/s41598-019-41090-8>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. <https://doi.org/10.1038/415429a>
- Feigin, H. (2021). Perceptual decisions are biased toward relevant prior choices. *Scientific Reports*, *16*.
- Fischer, C., Czoschke, S., Peters, B., Rahm, B., Kaiser, J., & Bledowski, C. (2020). Context information supports serial dependence of multiple visual objects across memory episodes. *Nature Communications*, *11*(1), 1932. <https://doi.org/10.1038/s41467-020-15874-w>
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. <https://doi.org/10.1038/nn.3689>
- Fornaciai, M., & Park, J. (2018). Attractive Serial Dependence in the Absence of an Explicit Task. *Psychological Science*, *29*(3), 437–446. <https://doi.org/10.1177/0956797617737385>
- Fornaciai, M., & Park, J. (2020). Attractive serial dependence between memorized stimuli. *Cognition*, *200*, 104250. <https://doi.org/10.1016/j.cognition.2020.104250>
- Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite Effects of Recent History on Perception and Decision. *Current Biology*, *27*(4), 590–595. <https://doi.org/10.1016/j.cub.2017.01.006>
- Fritsche, M., Spaak, E., & de Lange, F. P. (2020). A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *eLife*, *9*, e55389. <https://doi.org/10.7554/eLife.55389>
- King, J.-R., & Wyart, V. (2021). The Human Brain Encodes a Chronicle of Visual Events at each Instant of Time thanks to the Multiplexing of Traveling Waves. *The Journal of Neuroscience*, JN-RM-2098-20. <https://doi.org/10.1523/JNEUROSCI.2098-20.2021>
- Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. *Trends in Cognitive Sciences*, *21*(7), 493–497. <https://doi.org/10.1016/j.tics.2017.04.011>
- Kohn, A. (2007). Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. *Journal of Neurophysiology*, *97*(5), 3155–3164. <https://doi.org/10.1152/jn.00086.2007>
- Kondo, A., Takahashi, K., & Watanabe, K. (2012). Sequential Effects in Face-Attractiveness Judgment. *Perception*, *41*(1), 43–49. <https://doi.org/10.1068/p7116>
- Liberman, A., Fischer, J., & Whitney, D. (2014). Serial Dependence in the Perception of Faces. *Current Biology*, *24*(21), 2569–2574. <https://doi.org/10.1016/j.cub.2014.09.025>
- Liberman, A., Zhang, K., & Whitney, D. (2016). Serial dependence promotes object stability during occlusion. *Journal of Vision*, *16*(15), 16. <https://doi.org/10.1167/16.15.16>
- Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 369–380. <https://doi.org/10.1037/0278-7393.34.2.369>

- Manassi, M., Liberman, A., Chaney, W., & Whitney, D. (2017). The perceived stability of scenes: Serial dependence in ensemble representations. *Scientific Reports*, *7*(1), 1971. <https://doi.org/10.1038/s41598-017-02201-5>
- Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., & Whitney, D. (2018). Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review*, *25*(6), 2245–2253. <https://doi.org/10.3758/s13423-018-1454-5>
- Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods*, *128*(1–2), 85–93. [https://doi.org/10.1016/S0165-0270\(03\)00151-1](https://doi.org/10.1016/S0165-0270(03)00151-1)
- Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLOS Biology*, *17*(3), e3000144. <https://doi.org/10.1371/journal.pbio.3000144>
- Rafiei, M., Chetverikov, A., Hansmann-Roth, S., & Kristjánsson, Á. (2021). You see what you look for: Targets and distractors in visual search can cause opposing serial dependencies. *Journal of Vision*, *21*(10), 3. <https://doi.org/10.1167/jov.21.10.3>
- St. John-Saaltink, E., Kok, P., Lau, H. C., & de Lange, F. P. (2016). Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *The Journal of Neuroscience*, *36*(23), 6186–6192. <https://doi.org/10.1523/JNEUROSCI.4390-15.2016>
- Taubert, J., Van der Burg, E., & Alais, D. (2016). Love at second sight: Sequential dependence of facial attractiveness in an on-line dating paradigm. *Scientific Reports*, *6*(1), 22740. <https://doi.org/10.1038/srep22740>
- Tulving, E., & Schacter, D. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306. <https://doi.org/10.1126/science.2296719>
- van Bergen, R. S., & Jehee, J. F. M. (2019). Probabilistic Representation in Human Visual Cortex Reflects Uncertainty in Serial Decisions. *The Journal of Neuroscience*, *39*(41), 8164–8176. <https://doi.org/10.1523/JNEUROSCI.3212-18.2019>
- Van der Burg, E., Toet, A., Brouwer, A.-M., & van Erp, J. B. F. (2021). Sequential Effects in Odor Perception. *Chemosensory Perception*. <https://doi.org/10.1007/s12078-021-09290-7>
- Xia, Y., Leib, A. Y., & Whitney, D. (2016). Serial dependence in the perception of attractiveness. *Journal of Vision*, *16*(15), 28. <https://doi.org/10.1167/16.15.28>

APPENDIX



Appendix Figure 5: Statistical results pertaining to hypothesis tests in Fig. 2 as a function of swap error classification threshold. (upper) Permutation test p-values for different probability threshold of swap error exclusion. For example: a threshold of 0.2 indicates that all trials which were classified as swap errors with a probability of 0.2 or more were removed prior to data analysis. The black line indicates $\alpha = 0.05$. (lower) This panel indicates the amount of trials that were classified with a given probability of swap errors.



Appendix Figure 6: Error distribution histograms for each of the four orientations presented in Experiments 1, 2B and 3.

CHAPTER 3: DISCUSSION



3.1 What have we learned? A brief summary.

In “**Error compensation in random vector double step saccades with and without global adaptation**” I showed how object coordinates are updated across eye movements, even when targets are presented at random locations. Errors made during eye movement execution are accurately accounted for and this is true for random errors as well as errors induced by systematically changing saccade amplitudes through saccadic adaptation. This error correction is based on the actually executed, not the planned eye movements. My analysis approach provided a way of investigating the spatial remapping process across saccades of any amplitude, and towards any location, rather than exclusively the highly stereotyped saccade sequences employed in the field.

In “**Remapping high-capacity, pre-attentive, fragile sensory memory**” I showed that it is not only attended, robust working memory representations are updated across eye movements, but weakly attended (or unattended) objects that are not typically classified as working memory are also updated. This contradicts previous views, which stated that only attended objects could be remapped. Instead, it is more likely that many or even all perceived objects are remapped across eye movements, irrespective of whether they were in the focus of attention or not.

In “**The development of retro-cue benefits with extensive practice**” I showed that retrospectively cueing short-term memory representations is unlikely to reflect the process of bringing implicit memory representations (i.e., iconic or fragile memory) into stable and reportable working memory. Instead, retro-cues shift memory resources within the same memory system (working memory). As a result, participants are able to report even those representations that were initially weakly encoded and would not have survived the recall process without first being prioritized via the retro-cue. These findings also imply that estimates of working memory capacity represent the measurement of a construct defined by the employed paradigm, rather than being pure measure of a cognitive

capacity. The true capacity of human working memory is therefore still unknown.

In **“Successful visually guided eye movements following sight restoration after congenital cataracts”** I investigated basic visually-guided eye movements in individuals who were born blind (unable to perceive any patterned visual input) due to congenital bilateral total cataracts, and who received a sight-restoring cataract operation later in life. I revealed that these individuals were able to make goal-directed eye movements despite an absence of vision during the critical first year of their lives. They presented with severe clinical nystagmus (like most individuals with uncorrected early-life visual deficits) but were not worse at making eye movements towards visual targets than a control group with similar nystagmus, including saccades. This was good news for such cataract patients as eye movements are one of the most fundamental elements of human vision. My results imply that even without vision during the early-life sensitive phase, the saccadic eye movement system could still develop later in life, when vision was restored. In addition, I developed a method for calibrating eye trackers in any individual with nystagmus or other eye movement disorders that prevent stable target fixation. This algorithm opens up the way for research into visually guided eye movement dynamics in individuals with severe nystagmus, including sight-recovered cataract individuals. Research in this area has so mostly far been limited to the dynamics of the nystagmus waveform without mapping eye position to screen coordinates.

In **“Serial dependence reflects integration of task-relevant information in working memory”** I discovered that it is not only previously observed visual information that influences current perception and behavioural responses, but that all information concurrently represented in working memory influences behavioural responses to individual percepts. That is, responses were biased toward similar and task-relevant information (indicating integration) and biased away from dissimilar and task-irrelevant information (indicating repulsion). These results revealed the crucial role of working memory in visual history effects (such as serial dependence), a field of research that has received increased interest in recent years. We propose that, similar to multisensory perception, several observations of the same object could be integrated to increase the

reliability of an object representation based on a single instance of perception, in order to facilitate more effective object-oriented behaviour.

3.2 Samenvating in het Nederlands

Bij het maken van een oogbeweging landen de ogen niet altijd op de bedoelde plek. Tijdens het maken van de eerstvolgende oogbeweging kan het visuele systeem corrigeren voor eventuele fouten die tijdens de uitvoering van de eerste oogbeweging zijn gemaakt. In Hoofdstuk 1 (“Error compensation in random vector double step saccades with and without global adaptation”) laat ik zien dat deze foutcorrectie is gebaseerd op de daadwerkelijk uitgevoerde oogbeweging, en niet op de geplande oogbewegingen. Bestaande studies toonden dit eerder aan voor een zeer beperkt palet aan oogbewegingen, waarbij participanten herhaaldelijk dezelfde voorspelbare sequenties van oogbewegingen moesten maken. Een nieuw experimenteel design en analyseaanpak maakte het mogelijk om de correctie van fouten tijdens achtereenvolgende oogbewegingen te onderzoeken tijdens meer natuurgetrouw kijkgedrag met oogbewegingen van wisselende amplitudes en naar willekeurige locaties.

Wanneer we een groot aantal objecten gelijktijdig proberen te onthouden kunnen we er naderhand maar een paar van rapporteren. Hieruit wordt doorgaans geconcludeerd dat het menselijk kortetermijngeheugen een kleine capaciteit heeft van slechts (grofweg) vier objecten. Wanneer we echter op het laatste moment geïnstrueerd worden welke specifiek object we moeten rapporteren (aan de hand van een retrospectieve ‘cue’), blijken we veel meer objecten te kunnen rapporteren. Een interpretatie van deze bevinding is dat sommige objecten in het geheugen meer worden geattendeerd (deze objecten maken deel uit van het robuuste werkgeheugen) terwijl andere objecten niet of nauwelijks worden geattendeerd (deze objecten zijn gevoeliger voor interferentie, en worden aangeduid als onderdeel van het “fragiele geheugen”). In Hoofdstuk 2 (“Remapping high-capacity, pre-attentive, fragile sensory memory”) laat ik zien dat na een oogbeweging, behalve de objecten in het robuuste werkgeheugen ook de objecten in het fragiele geheugen worden meegenomen naar de nieuwe kijklocatie. Deze bevinding is dan ook in tegenspraak met eerdere opvattingen, die stelden dat alleen geattendeerde objecten kunnen worden meegenomen tijdens het maken van een oogbeweging. De huidige data suggereert dat een veel rijkere representatie van de

buitenwereld behouden blijft tijdens het maken van een oogbeweging dan oorspronkelijk werd gedacht.

In Hoofdstuk 3 (“The development of retro-cue benefits with extensive practice”) onderzoeken we de opslagmechanismen die onderliggend zijn aan geattendeerde en ongeattendeerde objecten in het kortetermijngeheugen. Een hypothese is dat er een aparte opslag bestaat met hoge capaciteit, vanwaar we (aan de hand van de retrospectieve cue) objecten kunnen overdragen naar het robuuste werkgeheugen (dat een kleinere capaciteit heeft). Deze aparte opslag zou het zogenaamde ‘fragiele geheugen’ kunnen zijn, maar ook het langetermijngeheugen. Een alternatieve hypothese is dat zowel de ongeattendeerde als de geattendeerde objecten onderdeel uitmaken van dezelfde opslag (het werkgeheugen), maar dat onevenredige verdeling van aandacht over deze objecten ervoor zorgt dat sommige objecten wel direct rapporteerbaar zijn en andere objecten niet. De resultaten van de studie zijn in lijn met deze tweede hypothese. Deze bevinding impliceert dat huidige meetmethodes van werkgeheugencapaciteit een onderschatting geven van het totaal aantal objecten in het werkgeheugen, omdat deze metingen alleen gebaseerd zijn op het aantal geattendeerde objecten; de ongeattendeerde objecten worden in deze inschattingen niet meegenomen. Bovendien laat de data zien dat zelfs de meting van het aantal geattendeerde objecten problematisch is: hoewel de werkgeheugencapaciteit een robuuste (bijna karaktertrek-achtige) maat zou moeten zijn, neemt de gemeten werkgeheugencapaciteit onafgebroken toe over herhaaldelijke metingen.

In Hoofdstuk 4 (“Successful visually guided eye movements following sight restoration after congenital cataracts”) onderzoek ik doelgerichte oogbewegingen bij personen die door cataracten aan beide ogen volledig blind zijn geboren, en voor het eerst konden zien na een staaroperatie op latere leeftijd. Ondanks een gebrek aan gezichtsvermogen tijdens het kritieke eerste jaar van hun leven, toonde de data aan dat deze personen alsnog in staat waren om doelgerichte oogbewegingen te maken. Ze vertoonden ernstige klinische nystagmus (een aandoening waarbij de ogen ongecontroleerde ritmische bewegingen maken), maar waren niet slechter in het maken van oogbewegingen naar visuele doelen dan een controlegroep met vergelijkbare nystagmus. Dit is een zeer

hoopgevende bevinding voor cataractpatiënten (en andere patiënten die blindgeboren zijn door een afwijking aan de ogen), aangezien succesvolle oogbewegingen één van de meest fundamentele aspecten van het menselijk gezichtsvermogen zijn. Tijdens de uitvoer van dit onderzoek heb ik een methode ontwikkeld voor het kalibreren van oogmeetapparatuur bij personen met nystagmus en andere oogbewegingsstoornissen die stabiele fixatie in de weg staan. Deze methode kan breed ingezet worden voor allerlei toekomstig onderzoek naar populaties met nystagmus, een onderzoeksveld dat tot dusverre grotendeels beperkt bleef tot het karakteriseren van de nystagmus zelf (in grootte en richting), zonder daarbij de oogpositie in kaart te kunnen brengen.

Doordat we continu oogbewegingen maken moet ons visuele systeem de waarneming van de wereld om ons heen opbouwen uit sequenties van losse ‘filmfragmentjes’. Het kortetermijngeheugen maakt het mogelijk om de verschillende filmfragmentjes over tijd aan elkaar te lijmen tot een coherent percept. De verschillende filmfragmentjes beïnvloeden elkaar echter ook. Een recente bevinding is dat als mensen een precieze visuele eigenschap van een object moeten rapporteren (bijvoorbeeld de kleur of oriëntatie), deze rapportage verschuift in de richting van een eerder waargenomen object. Bijvoorbeeld, als de kleur rood wordt gevolgd door de kleur blauw, zal de kleur blauw als roder worden gerapporteerd dan wanneer tweemaal dezelfde kleur blauw was getoond. Dit fenomeen heet “serial dependence”. In Hoofdstuk 5 (“Serial dependence reflects integration of task-relevant information in working memory”) demonstreer ik dat het niet alleen eerder waargenomen visuele informatie is die rapportage van de huidige visuele informatie beïnvloedt, maar dat deze beïnvloed wordt door alle informatie die op dat moment in het werkgeheugen is opgeslagen. Zo laat ik zien dat de oriëntatie van een visueel object zelfs (met terugwerkende kracht) wordt beïnvloed door de oriëntatie van een later getoond visueel object. Er zijn veel indicaties dat dit beïnvloedingsproces een adaptieve functie dient. Zo hebben objecten die op het doelobject lijken een aantrekkende kracht op de gerapporteerde waarneming, terwijl objecten die niet op het doelobject lijken juist een afstotende kracht hebben op de gerapporteerde waarneming. Bovendien worden deze verschuivingen in de gerapporteerde waarneming alleen veroorzaakt door objecten die relevant zijn voor de taak van de

proefpersoon (en dus geattendeerd worden). Deze resultaten onthullen de cruciale rol van het werkgeheugen in het combineren van losse observaties (of filmfragmentjes) over tijd. In meer fundamentele zin stellen we voor dat meerdere waarnemingen van hetzelfde object worden geïntegreerd in het werkgeheugen om zo de betrouwbaarheid van een objectrepresentatie te vergroten (ten opzichte van de objectrepresentatie voortkomend uit een enkele observatie), ten behoeve van effectief doelgericht gedrag.

3.3 The perpetual diaspora of early career researchers

It is an unwritten rule, bekownst to nearly all academic researchers who did not (yet) secure a tenured professorship, that one has to “go abroad” to “help one’s career”. The university where one executed hir undergraduate studies rarely becomes the place where one goes on to do graduate and post-graduate studies. And then, in the afterglow of a successfully defended PhD one is typically encouraged to go to a different country for a post-doc (that is, if one is indeed still hell-bent to pursue a career in academia). More often than not, there is no need for such encouragement as there are simply no positions available at one’s “home institution”, and to stay in academia, one has no choice but to move to a new city and typically even a new country and continent. One is told that this is simply how it works. If someone wants to pursue this career, hir cannot miss any opportunity. To some degree this makes sense. A different lab will offer different experiences, teach one different skills, and broadens one’s meaningful professional network. However, while this seems to be generally accepted as the way of things, it comes at a cost. Many senior wizards might have forgotten what it means to give up your life and move to a different country. It means to leave behind the life you built around a network of people, local knowledge, and foundational root structures. It means to leave behind friends, familiar routines, and survival strategies, in many cases even lovers and family. It means starting from scratch in a new city, as a stranger, possibly not speaking the local language. There may be a feeling of not knowing anybody, of nobody caring one bit about you, of having to re-grow one’s social network and practical local knowledge from absolutely nothing. It requires an incredible amount of resilience, of optimism, of tireless social skills and of mental health stability to attempt such endeavours. The worst part is that very likely this new home will have to be abandoned, again, after a short time, to follow through to the next opportunity, the next post-doc position, like chasing a purple dragon at the expense of one’s life.

As of this writing, since my undergraduate studies I have lived in 7 cities and 5 countries and am now about to move to Iceland to start my first, and likely not last, post-doc. I am certainly immensely grateful for the experiences I had in all of these places, the many opportunities I was given to grow as a scientist, and all the good friends I eventually made along the way. However, this life has taken a toll, and I am tired. Tired of repeatedly giving up everything for a science career. Tired of being in my 30s and not knowing where I will be in a few years, of not being able to buy a house and start a family. Tired of the perpetual job insecurity and the choice to either accept this fate, or to leave academia, as so many others have. Most former academics I talked to did not regret their decision to leave. This is a personal reflection, but I know that most, if not all, ECRs feel this. The situation is more tragic than just the fate of these individuals. It is tragic for academic science itself. The so-called brain drain to industry is real. And who wouldn't want a job that is better paid, has regulated work flows and conditions, where one can focus on excelling on a specific set of skills (rather than the requirement to do everything from teaching to grant writing to statistics to graphics design to management), where a work-life balance is often actively encouraged and where one is not haunted by time-limited contracts, or at least not as much. That is not to say that industry jobs are a heaven on earth and better in every regard (see the last section of this discussion for a positive light on academia), but commercial businesses do attract talented people away from academia for clear reasons. Industry jobs may not last forever either, but at least one would be likely to find the next job within commuting distance, while in academia, the next lab doing related research may be thousands of kilometres away. Academic science needs to step up its game if the frontier of human knowledge is to be advanced continuously and with the involvement of the smartest and most capable people.

These thoughts were recently reflected in a movement among German scientists, known under the hashtag #IchBinHannah. One hopeful legislative move came from Berlin, where the Hochschulgesetz (“university law”; see e.g., <https://ichbinhanna.wordpress.com/english-version/>) was recently adapted to ensure that new PhD candidates would find permanent positions, be that in the form of tenured professorships or researcher positions resembling post-docs but with permanent employment

contracts. This is precisely the kind of move many post-docs would like to see, including myself. Of course, the response of universities was to pause all hiring of post-docs until the state would actually free the resources necessary to implement this ambitious plan.

While I would like to do the full professor routine, what I really want to do is research, hands-on research in the lab, in a small team of brilliant scientists. I don't necessarily want to be in, or lead a mega-lab, I don't want to manage 15 PhDs, I want to do science: robust, reproducible, meaningful, fundamental science, using the skillsets I have acquired over the years, and continue to learn more. Ironically, many of the professors I have talked to now spend only a fraction of their time doing actual research, with the remaining time being taken up by admin work, managing their subordinates, university politics and teaching. In the Netherlands and many other countries, the business model of a university is rooted in teaching, that is, student tuitions and conditional government funding, which means that professors are often hired to do e.g., 60% teaching and 40% research. I don't mind teaching; I actually rather enjoy it. However, in the first years this means that in reality 90-100% of contractual time is spent doing teaching and class preparation (at least if those professors want their teaching to be any good at all), which leaves little time for research, most of which is spent writing grants. To further their career, emerging professors are forced to work overtime, leading to burn-out in many cases. The majority of the research work is thus left to their PhDs and post-docs, which makes it doubly troubling that these individuals are in a constant state of flux, rarely having the time to go deep and develop a longer line of research, which is crucial to develop meaningful theories based on reproducible results, something my field (cognitive science) continues to struggle with.

One possibility that has rarely been discussed and that started to appear more during COVID times is the idea of remote work. At least in some fields (such as mine), it crystallized that many, if not most experiments don't need a dedicated lab environment and in-person data collection, and that online experiments deliver high-quality data. I am able to collect hundreds of participant hours' worth of data in just one day, saving a great deal of experimenter time, where previously it took weeks to collect data from a small number of participants. This turn of events led me to an interesting thought: what

if two PI's from different universities (e.g., a former and future supervisor) team up and co-supervise a post-doc. For example, let's say that Dr. Feynman just completed her 2-year post-doc contract at Amsterdam University. She got a new post-doc contract at Frankfurt university. In the old world, she would have to say good-bye to her Dutch friends and move to Germany. What if she stayed in Amsterdam, continued to collaborate with her former supervisor and continued to benefit from the facilities (e.g., EEG labs) and social networks (e.g., lab meetings) of Amsterdam University, while being funded by Frankfurt University with her new PI as senior author on her publications. Not only does this allow Dr. Feynman to continue her life in Amsterdam with occasional trips to Frankfurt, but it also fosters collaborations within a larger academic network. Obviously, there are many potential bureaucratic hurdles to overcome for this scenario to blossom, and this would work better in a within-EU or within-US context, but in principle, and if all parties agree, this should be an entirely conceivable scenario. The idea was recently picked up and discussed by one of my favourite scientific podcasts, Everything Hertz (<https://everythinghertz.com/132>).

A change in the structure of academia, such as spearheaded for example by Berlin, is desperately needed (when in doubt, speak to any post-doc), and I can only hope that the senior scientists and university business executives take steps to ensure that all academic scientists, and especially ECRs, experience academia as the noble pursuit of knowledge that it can be under liveable conditions.

3.4 Fundamental vs. applied science

One subject I have discussed frequently with other researchers is fundamental science. I have heard from multiple sources, that this area of science is increasingly under threat in an ever more “business-like” university atmosphere. Fundamental science pertains to the kind of research that has no direct application in mind. At face value, the arguments against it make sense. Why should we spend taxpayer money on something that has no payoff in view? I would like to argue that it is absolutely crucial to maintain and even increase funding of non-applied science, in every country and every university. Furthermore, I believe that when it comes to writing grants, no promises of future applications should be required. Too often have I read statements of how, for example, a particular investigation into working memory will help people with this or that disorder, or how particular advances will inform clinical practice or even public policy. These statements are included either due to formal requirements or because researchers know who they face in grant and hiring committees, namely reviewers who would want to see this sort of thing.

Whether my work will have an impact on e.g., clinical practice or not, this should not be its primary motivation. Rather, to understand the world we live in is absolutely sufficient reason to invest in research into its fundamental, underlying substrates and mechanisms. It is precisely the absence of an applied goal that makes fundamental science so powerful and important. Without the constraint of an application to make money or cure people it is free to unravel the mysteries of the universe and the human mind and in doing so establish the foundation of understanding upon which everything else is built. Einstein, Maxwell, Kepler, and Curie did not need an application, but were simply driven by a burning curiosity to understand the physical world and its laws. Today, our civilization rests on the shoulders of these giants and their discoveries, while much of their research to this day remains theoretical, unapplied, waiting for the time that engineering and other areas of science have progressed to a point where e.g., relativistic effects start to matter a great deal more than they do today. Similarly, everything we do relies on that

marvellous and mysterious structure in our skulls and to this day we have barely begun to understand how it works. It is thus an area of science in which it is absolutely necessary to protect the freedom of scientists to research that which is there, simply because it exists and not because there is a visible societal benefit.

That all being said, I should make two disclaimers. First, I have the utmost respect for more applied scientists and am in no way suggesting that applied science is less important. To cure human diseases, to understand the spread of misinformation, to combat climate change in the here and now, to build the systems that make us go boldly into the future, all this is an absolutely noble, necessary, and important venture. My point here is mainly, that fundamental science *also* is necessary, and in fact, no applied science would be possible without its foundational roots. Second, I am not making a point for doing “useless” science or to distribute science funding arbitrarily. Great care must be taken to select the most important research areas, determined not by their currently perceived “sexiness”, or by their projected immediate payoff, but by how feasible and timely a project is given the current level of understanding and technology. In other words: given what has now been robustly established, what is the next aspect we need to understand? Too often are there brief trends in science, buzzwords ringing with the promise of impact factors and grants, often a re-invention of something from decades past, only to fade away again after a few years, without having established a strong theory, a strong foundation for the next step. Instead, I would like to see grant committees to fund slow science projects, aiming at digging deep and carving out theory that will not be overthrown again in a few years.

In short, I am suggesting that fundamental science always pays off, is always useful to society in the long run and as a whole, but that the actual application and use cases cannot and should not be determined prior to the actual quest of understanding the world.

3.5 Why then, would a rational agent choose this path?

Despite the critical thoughts in the previous section, I am still here, now beginning my first post-doc position and I am still in love with academic science as a profession. It is my dream job and I hope to remain for many years to come. While any area of human existence has its problems and perpetual miseries to improve, academic science to me remains one of the safe harbours of sanity and a noble path of truth seeking.

A big reason for this are the many people who are sailing with me in the same boat. In my field have very frequently encountered the most beautiful creatures. Driven by curiosity and a passion for learning and understanding, many people found their way into academic science, and not by coincidence, at the intersection between nature, mathematics, and art. It attracts all sorts, but most of us know why we are here: the freedom to follow our burning curiosity and to unleash our uniqueness into the generation of knowledge. In those moments when we are not gripped by imposter syndrome, we intuitively understand that there is only one person in existence who could do exactly the research we are doing. Being “different” is a boon in science, as the ability to think outside the box is absolutely fundamental to this endeavour. This, paired with largely an absence of profit-motives, brings together unique, quirky, highly intelligent, passionate, and compassionate individuals, who aren’t satisfied with driving in established comfortable grooves, but who make their own grooves, to boldly go where no woman has gone before, and to leave a mark in the world, which, in time, will likely benefit society in a small or large way.

It is this freedom that constitutes the second major point for why I still love academic science. At least in my current position, I am able to choose what I want to research. Within certain boundaries, it is not work that I pursue for someone else, not even for my supervisor (though they do have some say in it). The research I do is the result of recognizing what is known and to then go beyond that, into the unknown, to figure out one more tiny piece of the puzzle. I am the only

person in the world with my exact set domain-specific knowledge, and, as such, have some idea about where it would be most fruitful to go next. In turn, my work is not some assignment I deliver to someone else. It is my baby, and I care about it immensely. The energy obtained from this reality can be poured into even the most slow-paced, painstaking data analysis sessions and mind-bending thought trains.

Finally, there is a sense of purpose, beyond just monetary gains. I feel very privileged to work at the frontier of human knowledge, to struggle to push it just a little bit further. While industry jobs offer higher salaries and a more structured work environment and certainly can be very rewarding and interesting, the aim of academic science is to generate knowledge. This is where I want to be, where I found myself and what I identify with.

ACKNOWLEDGEMENTS

Nobody is successful alone.

First of all, I thank my parents, Helene and Hartmut, who were with me from the start and did everything and more in their power to support me, and who in my youth instilled in me a great love for nature, science, and art. You took me to museums and forests on a regular basis and you answered my endless questions about the world. You got me through school and university despite all obstacles and sacrificed a lot for me. No amount of thanks will ever be sufficient for you.

Vor allen anderen danke ich meinen Eltern, Helene und Hartmut. Ihr habt alles und mehr getan um mich in meinem Weg zu unterstützen und habt in mir von kleinauf die Liebe zur Natur, zur Wissenschaft und zur Kunst erweckt. Ihr seid mit mir regelmäßig in Museen und Wälder gegangen und habt meine endlosen Fragen über die Welt beantwortet. Ihr habt mich wider allen Hindernissen durch Schule und Universität gebracht und viel für mich aufgegeben. Es ist unmöglich euch ausreichend zu danken.

I thank my promotor Stefan for recognizing my potential when I was a wee master student and for patiently putting up with my stubbornness and quirks. Your remarkable ability to navigate the politics of academia enabled this PhD and you were always available when I needed help on a scientific, academic, or personal matter. You showed great understanding for my struggles and pushed me with just the right amount of force towards success, neither constraining me too much, nor letting me get lost in perfectionism purgatory. Without you I would still be trying to improve my first paper.

I thank my co-promotor Surya for the relentless optimism and extensive technical and scientific assistance as well as overall humaneness. Your hand-holding sessions guiding me through difficult analyses and offering reflections on my thinking, as well as elaborate feedback on my writing, were fundamental to my success. I met you as fellow PhD candidate, and while you obtained a well-deserved faculty

ACKNOWLEDGEMENTS

position in no-time, you never once made me feel inferior, but relentlessly and without judgment pushed me further towards success.

I thank Jutta, with whom I spent wonderful years, for accelerating my progress as a scientist and my productivity through lived example.

I thank Alexandra, with whom I spent wonderful years, for opening me up to myself and to the world of art.

I thank the secretaries of the departments I existed in, Eveline, Angelika, Janet, as well as the department managers, Ria and Eric and all other technical and non-technical support staff that worked behind the scenes. I also specifically thank all cleaning personnel and everyone else who might be overlooked but without whom civilization would not be possible. Your relentless efforts to support us researchers cannot be separated from our success as we would not be able to handle the administrative and other burdens without you.

I thank all my colleagues with whom I shared a work space, who offered advice and help, for no reason other than wanting to help a fellow human and scientist, and who were there to laugh, to complain and to understand. Jasper and Martijn, as office mates you marked the beginning of my scientific socialization and allowed me to see the joy of collaboration as well as learning together and teaching each other. Ada and Miriam, my time in Utrecht wouldn't have been the same without our regular chatting breaks and your advice and listening ear. Thank you Ignace for many great discussions in which you did not give me the answers to my questions but instead showed me how to improve my thinking process in order to start asking the right questions. Thank you Albert for believing in my potential early on and for opening the door. Thank you Gijs for wonderful scientific and non-scientific discussions and hangouts. Thank you Jan F. for inspiring me to pursue science. Thank you Sjoerd for your solid collegiality. Thank you Jim for your positive spirit. Thank you Yun for your early warning. Thank you Jessica for your inspiring superwoman-ness. Thank you Teuni for your calm, professional and rational approach. Thank you Siarhei for taking the time to sit with me and think. Thank you Chris P. for your caring and involvement. Thank you Chris J. for beaming positive energy into the world. Thank you Leon, for your lived example of accomplished professionalism. Thank you Marnix, for your inquisitiveness. Thank you Jose for your patience, understanding and

helpful collaboration. Thank you Suddha, for sharing your technical knowledge. Thank you Dennis S. for great discussions and for satisfying my electrophysiological urges. Thank you Ben for your insightful thoughts! Thank you Maria, you were a delightful presence and a ray of light during my time in Hamburg. Our lunch meetings were fundamental to my survival.

I thank Brigitte and Dr. Balu for giving me the opportunity to work with a special population. I thank Ramesh, Idris, Kabilan and Siddhart for supporting me in this endeavor. I thank the staff of the LVPEI for their relentless heroic deeds. I thank Sowmya and Rupesh for being my friends during the many months of data collection in Hyderabad.

I thank the many students whose thesis research projects I supervised. You inspired me to think about concepts clearly enough so that they could be taught, which in turn taught me as well.

I thank my teachers across 18 years of school for their enormous patience and in particular I would like to thank Frau Stein, Herr Kühn, Herr Schmälzle, and Herr Gröner for their inspirational way of teaching and nurturing my fascination for science.

Thank you Max and Charl, for being absolutely spectacularly supportive and wonderful friends through all these years and for serving as my paranymphs. Even though you have never met, both of you kept my fascination with how things work alive and did so much for me. Thank you Ioannis and Judith for being with during some difficult and formative years, for providing friendship, couch and calm ears, eyes and mouths. Furthermore, I thank all my friends, family and acquaintances who listened to and supported me, made me laugh and made me feel loved during these years. In particular (but no particular order) I thank Annika, Jonas, Michi, Andi, Amandine, Sneffy, Manni, Lydia, Balazs, Elvira, Lilia, Thomas, Lisa, Rob, Missy, Jelle, Fifi, Alex, Sofie, Jonte, Mike, Lily, Ben, Uzi, Giuseppe, Antonio, Casper, Tim, Anastea, Agatha, Moha, Fons, Jori, Echomaize.

There are certainly a number of people I could not remember but who helped me on my way. Don't take it personally, I will remember you the moment this goes to the printer. Thank you all!