Looking for Language in Space:

Spatial Simulations in Memory for Language

by

Alper Kumcu

A thesis submitted to the University of Birmingham
for the degree of DOCTOR OF PHILOSOPHY

July 2018
School of Psychology
College of Life and Environmental Sciences
University of Birmingham

Abstract

Grounded-embodied theories hold that language is understood and remembered through perceptual and motor simulations (i.e., activations and re-activations of sensorimotor experiences). This thesis aims to illustrate simulations of space in memory for language. In four experiments, we explored (1) how individuals encode and re-activate word locations and (2) how word meanings activate locations in space (e.g., “bird” - upward location). In the first part of the thesis (Experiment 1 and 2), we addressed the potential simulation of word locations by probing eye movements during memory retrieval (i.e., “looking at nothing”). In particular, we investigated why and when individuals need to rely on external memory support via simulation of word locations. Experiment 1 results reveal that the propensity to refer to the environment during retrieval correlates with individual’s visuospatial memory capacity. That is, participants with worse visuospatial memory relied more on the environment; whereas, participants with better visuospatial memory relied more on the internal memory sources. Experiment 2 shows that words which are more difficult to remember and, particularly, words that are more difficult to visualise in mind lead to more reliance on the environment during word retrieval. Experiment 1 and 2 suggest that the opportunistic and efficient human mind switches between internal sources and external support as a function of the richness of internal sources and cognitive demands coming from the words to be remembered. The second part of the thesis (Experiment 3 and 4), focuses on spatial simulations triggered directly by words (i.e., language-based simulations). Experiment 3 is a norming study in which raters were asked to associate words with locations in space. Experiment 3 results demonstrate that there is a high degree of agreement among individuals when linking both concrete and abstract words to locations in space although there are no explicit conventions with regard to these associations. Ratings in Experiment 3 indicate potential locations of word-induced simulations. Normed words were used as stimuli in Experiment 4 in which recognition memory for words with

spatial associations was probed. Experiment 4 results show that both language-based simulation of space and simulation of word locations dictate memory performance even if space is irrelevant and unnecessary for successful retrieval. In particular, words that were presented in incongruent locations as to the locations they imply (e.g., “bird” in a downward location) were remembered faster than words presented in congruent locations (e.g., “bird” in an upward location). Memory performance deteriorated whenever attention was shifted to the locations simulated with word meanings. Overall, the thesis specifies the mechanics of two different types of spatial simulation in language and their effects on memory. Results and their implications are discussed within the framework of grounded-embodied approaches to language and memory and the extended cognition.

Acknowledgement

What a roller-coaster of emotions! The thesis you are reading now was made possible through scientific, mental and emotional support of many people, which makes me a very lucky researcher. Here, I would like to name a few, and I apologise in advance to all others that I forgot to mention.

First and foremost, I would like to thank my supervisor Dr Robin L. Thompson for her never-ending support, meticulous supervision and invaluable insights. She shared my joy during success and she lifted me up during moments of anxiety (which I had a lot). Without her, I could not even imagine realising my dream of doing a PhD in psychology. Thank you for being a great person and a great supervisor.

Language and Psychology Lab Group had a huge impact on me as a researcher and the present thesis. I would like to thank Professor Linda Wheeldon for never leaving my questions unanswered, Dr Steven Frisson for his contributions to the present work with his questions and constructive feedback, Dr Katrien Segaert for her support in mixed-effects modelling and for being a role model with her research and lively presentations, Dr Andrew Olson, Dr Andrea Krott and also Professor Sotaro Kita from University of Warwick for their helpful suggestions on the experiments reported in this thesis.

My friends and colleagues here in UK deserve special thanks. Many thanks to Evi Argyriou for providing an inspiring example with her minimalist approach to research and presentation, Freya Atkins for being such a great (lab) mate, Mahmoud Elsherif for our fruitful debates about science and life in general, Isabella Fritz for our discussions in grounded-embodied cognition and R, Anneline Huck from University College London and Beinan Zhou for their support in collecting the pilot data, Linde Luijnenburg for her company at the British Library, Rachel Merchant for her support in recording the auditory stimuli, Kristian Suen for the teamwork in organising Language and Cognition Seminars (LanCoS) and Onur Usta for the library sessions and gallons of coffee.

I would like to acknowledge Turkish Council of Higher Education and Hacettepe University for awarding me with a doctoral scholarship and funding my research. I am grateful to my colleagues and friends at Hacettepe University in Turkey; especially Dr Sinem Sancaktaroğlu Bozkurt, Dr Hilal Erkazancı Durmuş, Dr Elif Ersözlü and Dr Sezen Ergin Zengin for smoothing out the funding process and for their lovely companionship.

I would like to thank all the participants who kindly participated the experiments. It goes without saying that the research could not be conducted without them. I also would like to thank our omniscient post-grad administrator Parveen Chahal for answering millions of questions that I have asked during my years at the School of Psychology.

Last but not least, I am deeply grateful to my mother and sister, who have sent their affection and support from miles away. The majority of the people I mentioned here is woman. A heartfelt appreciation to all women in science!

To my late father. You were and always will be a great scholar.

Table of Contents

List of Figures

List of Tables

List of Papers

The list below shows the papers published and/or accepted for publication.

Chapter 3 (Experiment 1)

Kumcu, A., & Thompson, R. L. (2016). Spatial Interference and Individual Differences in Looking at Nothing for Verbal Memory. In J. C. Papafragou, A., Grodner, D., Mirman, D., & Trueswell (Ed.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2387-2392). Austin, TX: Cognitive Science Society.

Chapter 4 (Experiment 2)

Kumcu, A., & Thompson, R. L. (2018). Less imageable words lead to more looks to blank locations during memory retrieval. *Psychological Research*. 1-18. https://doi.org/10.1007/s00426-018-1084-6.

Chapter 1

Introduction

1.1 Statement of Research Question

The link between space, language and memory is one of the most intriguing topics is cognitive psychology. Decades of evidence has showed that language provides us with a framework to materialise and structure the relatively abstract notion of space (Bloom, Garrett, Nadel, & Peterson, 1996; Carlson-Radvansky & Logan, 1997; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Miller & Johnson-Laird, 1976; Talmy, 1983). Thus, language influences how people perceive, think about and remember space (Levinson, 2003). The blooming fields of grounded-embodied cognition and the extended mind thesis have redefined the space-language-memory network and rekindled an interest about space with a novel perspective. Once seen as a central processing machine contained in the skull, the mind is now viewed as more of an interactive architecture extending onto the body and space. In accordance, burgeoning evidence suggests that space is not only a content but also a “medium” for language and memory (Mix, Smith, & Gasser, 2010). Memories can be indexed, and abstract thoughts can be grounded in space. Further, comprehending language can give rise to non-linguistic, spatial experiences. Despite the abundance of demonstrative experiments, how spatial perception and cognition influence language and memory operations is yet to be defined. To this end, the present thesis explores robust, systematic and often surprising ways that space is involved in memory for language. In particular, this thesis is an attempt to define and systematise different characteristics of spatial engagements during retrieval of words from memory. Thereby, it aims to contribute to the understanding of the effect of space on memory for language.

1.2 Definition of Concepts

The key concepts examined in the present work are defined in this section. Thereby, the scope of the thesis is outlined.

1.2.1 Spatial indexing

The human mind can anchor spatially-located information to external spatial locations. Simply put, the location of a visual item is encoded with the item itself. This behaviour is called as *spatial indexing*. Marr (1982) was among the first to argue that the visual system separates locations from visual features (what vs. where). He introduced the term *place token* to refer to representations of location. A place token indexes the locations of visual information at the early stages of visual processing. Pylyshyn (1989) operationalised the phenomenon of spatial indexing in an exhaustive model termed FINST. The model assumes that spatial indexing is a primitive, that is pre-attentive or a pre-cognitive mechanism, which precedes "higher" visual operations such as recognition of patterns. But importantly, spatial indices allow for stability of visual information by constructing stable representations of locations in a constantly changing visual world. Hence, an index keeps pointing to the same location even if the visual pattern moves across the retina. According to the model, therefore, spatial indexing is different from merely encoding the position of a feature because a spatial index makes it possible to locate the visual stimulus for further examination when the necessity arises (see also Ballard, Hayhoe, Pook, & Rao, 1997). In a similar vein, Coslett (1999) proposes a *spatial registration hypothesis*. This hypothesis assumes that "all stimuli, even if not relevant to the task at hand, are automatically marked with respect to spatial location in egocentric coordinate systems" (p. 703). Registering a location entails the creation of a marker that specifies the coordinate of an object in relation with the other objects in the environment. According to spatial registering hypothesis, spatial indexing is limited with the capacity of visual attention.

1.2.2 Looking at nothing

Consider the following situation: You are in the middle of a maths exam, trying to solve a problem. The problem you engage requires the application of a formula that you fail to remember at that point. However, you remember that the instructor has previously used the blackboard to explain the formula in one of the classes. You remember that she has written the formula on the top left corner of the board. Of course, the blackboard has already been cleaned and is totally blank now. Still, you raise your head in the hope of a sudden recall and look at the previous, but now-blank location of the formula. You have just looked at "nothing" (Spivey & Geng, 2000). Obviously, there is nothing on the board. What you are looking at is the spatial indice that represents the previous but now-blank location of the formula. In this respect, looking at nothing is a memory-guided behaviour. Your eyes have been oriented to the previous location of the formula because you have attempted to remember it. To be more precise, spatial indices tied to external visual and verbal information trigger eye movements when a mental representation is reactivated. Thus, when retrieving information from memory, people tend to exploit location-based indices and look at the seemingly uninformative, empty locations where the information originally occurred even if the location is irrelevant to the task. Looking at nothing follows spatial indexing in a typical memory task with encoding and retrieval stages. Individuals are expected to index the location of the information at encoding. Then, they are expected to look at the previous locations of the to-be-retrieved information during retrieval. A considerable number of empirical studies has documented looking at nothing (e.g., Altmann, 2004; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2000). The link between mental representations and looking behaviour (de Groot, Huettig, & Olivers, 2016; Ferreira, Apel, & Henderson, 2008; Martarelli, Chiquet, Laeng, & Mast, 2017; O'Regan, 1992; Renkewitz & Jahn, 2012; D. C. Richardson, Altmann, Spivey, & Hoover, 2009; D. C. Richardson & Spivey, 2000; Scholz, Mehlhorn, & Krems, 2011; Wantz, Martarelli, & Mast, 2015) and whether looks to blank locations improve memory (Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Johansson & Johansson, 2014; Scholz, Klichowicz, & Krems, 2018; Scholz, Mehlhorn, & Krems, 2016) have received much attention in looking at nothing research.

1.2.3 Simulation

In grounded-embodied cognition, a *simulation* is defined as a partial activation or reactivation of an original perceptual, motor, affective or introspective experience (Barsalou, 1999). A simulation can occur in the absence (offline) or upon the perception (online) of the original stimulus. A large body of neural and behavioural evidence indicates that major segments of cognition such as mental imagery, memory, language comprehension, consciousness, expertise and several cognitive performances such as facial mimicry, gesturing, reasoning and problem solving rely on simulations (see Dijkstra & Post, 2015; Hesslow, 2011; Körner, Topolinski, & Strack, 2016; Wilson, 2002 for reviews). The current thesis focuses on offline and online simulations of space in memory for language.

Perceptual symbols lie at the heart of simulation mechanism as they make it possible the (re)activation of sensorimotor experiences (Barsalou, 1999). Perceptual symbols are mental representations that represent conceptual knowledge in the mind. Crucially, perceptual symbols are represented in the same system as the perceptual states that produced them (Barsalou, 1999, p. 578). As a result, they reflect physical and thus, perceptual characteristics of the referents they stand for, and represent continuous, rich and multimodal phenomenological experience. That said, perceptual symbols "partly" represent their referents rather than being similar to “high-resolution video-clips” or “high-fidelity sound clips” (Zwaan, Stanfield, & Yaxley, 2002). Based on these features, perceptual symbols are fundamentally different from *physical symbols* (Newell, 1980) asserted within the computational theories of mind (Fodor, 1975; Haugeland, 1985; Newell & Simon, 1976; Putnam, 1960). In contrast to perceptual symbols, physical symbols are amodal, abstract and discrete units. That is, they do not reflect the perceptual modality of or do not resemble to the physical entities that they refer to as in 1 and 0’s in a computation environment (Harnad, 1990; Pylyshyn, 1986).

Offline simulation

Consider that you need to remember a piece of conceptual information about cats (e.g., “What does a cat sound like?”). According to grounded-embodied cognition, in such a case, the human mind relies on the reactivation of perceptual, motor, affective or introspective experiences that are formed during previous interactions with a cat (e.g., how soft its fur is, how it smells, how you feel when you stroke it etc.). However, these experiences or states are not reinstated exactly on later occasions and “different contexts may distort activations of the original representations” (Barsalou, 1999, p. 584). Offline simulations have a *situated* character (Barsalou, 2003). For instance, they represent specific cats in specific situations rather than representing generic knowledge about cats. You might have noticed that remembering a piece of information about cats in the way described above bear similarities with remembering the location of a maths formula. In the latter case, the individual encodes conceptual information (the formula) along with a spatio-perceptual experience (perceiving the location of the formula on the blackboard). Later, when she needs to access the conceptual information through memory retrieval, she reactivates the perceptual experience associated with it. In other words, she “re-lives” the perceptual experience that she has during the encoding of the conceptual information. On this ground, looking at nothing can be understood as a reflection of spatial simulation that takes place in the absence of the original stimulus.

Online simulation

Perceptual, motor, affective or introspective experiences can be activated whenever an individual perceives a stimulus. For example, merely viewing a graspable object such as a cup, a knife or a frying pan simulates the potential act of grasping. In turn, brain regions associated with motor movements are activated (Chao & Martin, 2000; Tucker & Ellis, 1998). Similarly, remembering a Japanese kanji character stimulates motor activity in the areas that would be activated when actually writing the characters (Kato et al., 1999; Topolinski & Strack, 2009). Merely listening to words that involve strong tongue movements when pronounced such as “birra” (beer in Italian) or “ferro” (iron in Italian) activates tongue muscles (Fadiga, Craighero, Buccino, & Rizzolatti, 2002). Language comprehension gives rise to simulations in this manner (see Chapter 2.1.3).

To summarise, the fundamental difference between an offline and an online simulation is the existence of an external stimulus at the time of the simulation. An offline simulation is a recreation of previous sensorimotor activations without any stimulus when the stimulus is re-accessed. An online simulation is a sensorimotor activation upon the perception of a stimulus (see Chapter 2.1 for further clarification with examples).

There are other instances of a simulation mechanism in cognitive psychology. In social cognition, for instance, attributing mental states to others as in “mind reading” (i.e., theory of mind) (Gallese & Goldman, 1998; Premack & Woodruff, 1978) is thought to be a simulation based on mirror neurons that are activated merely by observing others (Caggiano et al., 1996; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). This aspect of simulations is beyond the scope of the current thesis despite the possible overlaps between theory of mind and mental simulations (Shanton & Goldman, 2010). Likewise, simulation within the context of this thesis does not refer to computer-based simulations.

1.2.4 Cognitive offloading

Cognitive agents can perform physical actions to outsource their cognitive work to the body and the world with the aim of reducing cognitive load (Wilson, 2002). This behaviour is called as *cognitive offloading* (Risko & Gilbert, 2016). Making a shopping list instead of keeping items to buy in your mind is a typical example of such behaviour. Cognitive offloading is treated as a fundamental mechanism of *extended cognition*, which posits that there is not a strict border between the mind and the world with the mind "leaking" into the world in surprising ways (A. Clark, 1998; A. Clark & Chalmers, 1998).

Cognition can be extended to different body parts during several cognitive problems such as using hands in finger-counting (Butterworth, 2005) and co-speech gestures[1] in communication (see Goldin-Meadow & Wagner, 2005; Pouw, de Nooijer, van Gog, Zwaan, & Paas, 2014 for reviews). Cognitive work can be offloaded directly onto the world as well. In particular, people can exploit their immediate space in an intelligent way to reduce their cognitive load. In such a case, space becomes a resource that must be managed, much like time or energy (Kirsh, 1995). For example, Kirsh and Maglio (1994) showed that in Tetris, participants solve perceptual and cognitive problems (e.g., judging whether a geometric piece can fit into a specific position) in the space by physically rotating the pieces on the screen rather than in their minds which would require mental rotation.

Cognitive offloading with eye movements is a special case in which both body (via eye movements) and environment are used at the same time. In a seminal study, for instance,

[1] Note that cognitive offloading is only one suggested use of co-speech gestures. Co-speech gestures are assumed to play other roles as well (Goldin-Meadow, 1999).

Hayhoe, Bensinger and Ballard (1998) asked participants to copy a pattern of coloured blocks on a computer screen. Participants had to use a mouse to drag scrambled blocks from the source window to an empty workspace window. The colour of blocks in the pattern was changed at different points during the task. Results showed that participants launched frequent fixations to the original pattern when they were copying it into the workspace window rather than keeping the original pattern in mind. In this respect, looking at nothing can be conceptualised as a type of cognitive offloading in which eyes are used to index certain locations in space for subsequent use in line with information-gathering goals (Ballard et al., 1997).

There are two important aspects of cognitive offloading. First, cognitive offloading often results in betterment in performance. For example, children were found to be most accurate when they actively gesture to count (even when using a puppet to count for them) as compared to internal counting (Alibali & Dirusso, 1999). Second, exploitation of body and space to reduce cognitive load is correlated with the cognitive demands coming from the task and internal, "biologic" cognitive capacity. That is, higher cognitive load and/or lower cognitive capacity result in more frequent cognitive offloading. For example, in Risko and Dunn (2015), participants were asked to remember a mixture of random letters in a traditional short-term memory task. They either had to rely on internal memory only or had the option to write down the presented information and thus, to externalise memory work. Results showed that use of external storage became more frequent as the number of letters in a string increased. Further, individual short-term memory capacity predicted the likelihood of writing down the to-be-remembered information. The functional role of cognitive offloading and the link between cognitive load/capacity and the tendency of offloading indicate that there is a systematic trade-off between internal and external processing (Schönpflug, 1986). In the face of costs of cognitive operations and limited cognitive sources, a successful cognitive agent seeks to "get the job done" in the easiest way possible by integrating internal with external processes in complex environments (A. Clark, 1989).

1.3 Overview of Thesis

The current thesis investigates activations and re-activations of space in memory for language. It contains seven chapters (including this one).

Chapter 2 includes the review of the literature. Chapter 2.1 focuses on the simulations in mental imagery, memory and language as the core mechanism of grounded-embodied cognition. Chapter 2.2 addresses the relation between eye movements and memory.

Chapter 3 contains the first experimental study in the form of a paper. This study investigates whether individuals with better visuospatial memory relies more on space and simulation of word locations through looks at previous but now-blank locations when retrieving words from memory.

Chapter 4 contains the second experimental study in the form of a paper. This study employs the experimental design developed in Chapter 3 to investigate whether words that are more difficult to maintain and retrieve from memory lead to more reliance on space and spatial simulations via looking at nothing. It also examines the contributions of word properties (e.g., word length, imageability, frequency etc.) to the looking behaviour and memory performance.

Chapter 5 contains the third experimental study in the form of an article. This is a lab-based norming study in which participants were asked to read 1439 concrete and abstract words and associate them with horizontal and vertical locations on a two-dimensional coordinate system.

Chapter 6 contains the forth experimental study in the form of a paper. This study uses spatially normed words from Chapter 5 to investigate how spatial locations suggested and simulated by word meanings affect recognition memory performance in relation with the physical locations of the words on the screen.

Finally, Chapter 7 (General Discussion) summarises findings and conclusions of the empirical studies and discusses them in the context of the grounded-embodied and extended approaches to memory and language. Additionally, it includes suggestions for future work.

Chapter 2

Theoretical Background

2.1 Mind, Recreated: Simulations in Imagery, Memory, and Language

2.1.1 Mental imagery as a simulation

Mental imagery is the ability to construct mental representations in the absence of external sensory stimulation. Thus, it is a *quasi-phenomenal* experience (N. J. T. Thomas, 1999, 2018a). That is, it resembles the actual perceptual experience but occurs when the appropriate external stimulus is not there. The ability to see with the "mind's eye" without any sensory stimulation is a remarkable feature of the human mind. Mental imagery underlies our ability to think, plan, re-analyse past events or even fantasise events that may never happen (Pearson & Kosslyn, 2013). Accordingly, mental images involve, alter or even replace the core operations of human cognition such as memory (Albers, Kok, Toni, Dijkerman, & De Lange, 2013; Rebecca Keogh & Pearson, 2011; Tong, 2013), problem-solving (Kozhevnikov, Motes, & Hegarty, 2007) decision-making (Tuan Pham, Meyvis, & Zhou, 2001), counter-factual thinking (Kulakova, Aichhorn, Schurz, Kronbichler, & Perner, 2013), reasoning (Hegarty, 2004; Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003), numerical cognition (Dehaene, Bossini, & Giraux, 1993) and creativity (LeBoutillier & Marks, 2003; Palmiero et al., 2016). The human mind can mentally "visualise" not only visual but also nonvisual perceptions (Lacey & Lawson, 2013) such as auditory mental imagery (e.g., imagining the voice of a friend or a song) (Lima et al., 2015) or motor mental imagery (e.g., mentally rehearsing a movement before actualising it) (Hanakawa, 2016). Mental images can arise from nonvisual modalities (particularly auditory or haptic) in congenitally blind individuals (Cattaneo et al., 2008). However, the literature of mental imagery is largely dedicated to metal imagery that is specifically visual (Tye, 1991).

How do we imagine? By extension, what does a mental image look like? The format of mental images has been extensively discussed in the 70s and 80s with two camps: pictorial (depictivism) and propositional (descriptivism) imagery. The pictorial position (Kosslyn, 1973) holds that mental images are like pictures and there are spatial relations between the imagined objects. On the other hand, the propositional view (Pylyshyn, 1973) is that mental images are more like linguistic descriptions of visual scenes based on *tacit knowledge* about the world (i.e., implicit knowledge that is difficult to express explicitly such as the ability to ride a bike). Mental imagery under the treatment of descriptivism is more of an amodal, formal system. Whereas, depictivism offers a picture of mental imagery that appears more compatible with the mechanics of grounded-embodied cognition. However, neither of these approaches captures the true essence of grounded-embodied cognition because both of them depend on an information processing approach. In both cases, perceptual data flows inward to a passive cognitive agent (N. J. T. Thomas, 1999).

On the other hand, grounded-embodied theories of cognition conceive that mental imagery is based on active perceptions and actions. Mental images are considered as mental representations reactivated through previous perceptions (Ballard et al., 1997). Consequently, mental imagery is a simulation itself (Barsalou, 1999). As a matter of fact, mental imagery is assumed to be the most typical example of the simulation mechanism in that there are certain similarities between the properties of a sensorimotor simulation and mental imagery (Markman, Klein, & Suhr, 2008): First, mental images arise from perceptual representations. They are formed in the absence of the original perceptual stimulation. And lastly, a mental image is not an exact copy of the percept but rather, a partial recreation (Kosslyn, 1980). Within this view, the primary function of mental imagery is to simulate reality "at will" in order to access previous knowledge and predict the future (i.e., *mental emulation*) (Moulton & Kosslyn, 2009).

In order to verify that mental imagery is sensorimotor simulation, evidence showing similarities between perception and imagery is needed. This is indeed what the literature on mental imagery within the framework of grounded-embodied cognition indicates. For instance, an overwhelming body of neuroimaging evidence shows that similar brain regions are activated during perception and imagery stages (Cichy, Heinzle, & Haynes, 2012; Ganis, Thompson, & Kosslyn, 2004; Ishai & Sagi, 1995; Kosslyn, Thompson, & Alpert, 1997; O'Craven & Kanwisher, 2000). Behavioural studies further reveal the nature of the link between perception and imagery. As early as 1910, the psychologist Cheves Perky's experiments showed that visual mental images can supress perceiving real visual targets unconsciously (i.e., *the Perky effect*) (Craver-Lemley & Reeves, 1992; Perky, 1910). In the original experiment, participants were asked to fixate a point on a white screen and visually imagine certain objects there such as a tomato, a book or a pencil etc. After a few trials, a real but a faint image (i.e., in soft focus) of the concerned object was projected onto the screen. Participants failed to distinguish between their imagined projections and the real percepts. Shortly, real images intermingled with the mental images. For instance, some participants reported their surprise when they "imagined" an upright banana rather than a horizontally oriented one they were attempting to imagine (N. J. T. Thomas, 2018b). The Perky effect indicates that mental imagery and visual perception draw on the same sources (see also Finke, 1980).

In a similar study (Lloyd-Jones & Vernon, 2003), participants saw a word (e.g., "dog") accompanied by a line drawing of that object in the perception phase. In the imagery phase, participants made spatial judgements about the previously shown picture. Simultaneously, a picture distractor appeared on the screen during mental imagery. The picture distractor was either unrelated to the mental image of the previously shown object (e.g., dog - strawberry) or conceptually related (e.g., dog - cat). Response times in the judgement task were longer when participants generated a mental picture along with the perception of a conceptually related picture but not a conceptually unrelated picture. These findings suggest that imagery and visual perception share the same semantic representations. Mental images are also processed in similar ways as the actual images. In Borst and Kosslyn (2008), participants scanned a pattern of dots and then, an arrow was shown on the screen. Participants then decided whether the arrow pointed at a location that had been previously occupied by one of the dots. Results showed that the time to scan during imagery increased linearly as the distance between the arrow and the dots increased in perception. Further, participants who were better at scanning distances perceptually were also better at scanning distances across a mental image, suggesting the functional role of perception in mental imagery.

Finally, eye movement studies have given considerable support to the simulation account of mental imagery with two key findings (see Laeng, Bloem, D'Ascenzo, & Tommasi, 2014 for a review): First, eye movements during perception are similar to those during imagery (Brandt & Stark, 1997; Johansson, Holsanova, & Holmqvist, 2006). Second, the amount of overlap between eye movements during perception and imagery predicts the performance in imagery-related tasks (Laeng & Teodorescu, 2002). Eye movements in mental imagery are further elaborated in Chapter 2.2.3.

2.1.2 Memory retrieval as a simulation

A simulation account of memory views memory retrieval as a partial recreation of the past that often includes sensorimotor and contextual details of the original episode (see Buckner & Wheeler, 2001; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; Danker & Anderson, 2010; Kent & Lamberts, 2008; Pasternak & Greenlee, 2005; Rugg, Johnson, Park, & Uncapher, 2008; Xue, 2018 for comprehensive reviews and see De Brigard, 2014; Mahr & Csibra, 2018; Marr, 1971 for theoretical discussions). Hence, memory retrieval can be thought as a simulation of encoding in a similar way to mental imagery being a simulation of perception.

Indeed, it is known that mental imagery and memory operate on similar machinery as long as their perceptual modalities match (e.g., visual mental imagery - visual memory). In the original model of working memory, Baddeley and Hitch (1974) assumed that one function of the *visuospatial sketchpad* (i.e., the component of working memory responsible for the manipulation of visual information) is manipulating visual mental images. In support of this assumption, Baddeley and Andrade (2000) showed that visual and auditory mental imagery tasks disrupted visual and auditory components of working memory respectively; that is, visuospatial sketchpad and *phonological loop* (i.e., the component of working memory responsible for the manipulation of auditory information). Keogh and Pearson (2014) evidenced that individuals with stronger visual mental imagery also have greater visual working memory capacity but not verbal memory capacity (see also Keogh & Pearson, 2011).

Grounded-embodied cognition takes the link between mental imagery and memory one step forward: Memory not only involves mental imagery, but memory is mental imagery itself. In accordance, encoding corresponds to perception and retrieval corresponds to imagery. In this respect, Albers et al. (2013) presented strong evidence that working memory and mental imagery share representations in the early visual cortex (V1 - V3). Further, as Buckner and Wheeler (2001) noted "assessments of visual mental imagery ability in patients with damage to visual cortex support the possibility that brain regions involved in perception are also used during imagery and remembering" (De Renzi & Spinnler, 1967; D. N. Levine, Warach, & Farah, 1985).

Mental time travel is a striking example of the role of imagery in memory (Corballis, 2009; Schacter, Addis, & Buckner, 2007; Suddendorf & Corballis, 2007; Szpunar, 2010). Mental time travel is a cognitive ability of episodic memory (i.e., conscious and explicit recollection of past events) and episodic future thinking through mental imagery. Thus, a mental time traveller can mentally project herself backwards in time to re-live (i.e., reconstruct) the past events and pre-live (i.e., predict) the possible future events (Suddendorf & Corballis, 2007). In this respect, mental time travel can be considered as an *intertemporal simulation* (Shanton & Goldman, 2010). Growing evidence has shown that episodic memory and simulation of future by mental imagery share a core neural network (i.e., default network) (see Schacter et al., 2012 for a review), suggesting that memory, mental imagery and thinking about future rest on the similar neural mechanisms.

As in mental imagery, a simulation approach to memory underlines the correspondence between encoding and retrieval (Kent & Lamberts, 2008). Mounting evidence illustrates that common neural systems are activated both in encoding and retrieval (Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler, Petersen, & Buckner, 2000). Crucially, the similarity between neural patterns during encoding and retrieval is often predictive of how well an experience is remembered subsequently (see Brewer, Zhao, Glover, & Gabrieli, 1998; Wagner et al., 1998 for reviews). There is much evidence indicating that reinstated neural activations are specific to perceptual modality (visual vs. auditory), domain (memory for what - where) and feature (colour, motion or spatial location) (see Slotnick, 2004 for a review). For example, Wheeler, Petersen and Buckner (2000) gave participants a set of picture and sound items to study and then a recall test during which participants vividly remembered these items. Results demonstrated that regions of auditory and visual cortex are activated differently during retrieval of sounds and pictures. In a similar fashion, Goldberg, Perfetti and Schneider (2006b) asked participants whether a concrete word possesses a property from one of four sensory modalities as colour (e.g. green), sound (e.g., loud), touch (e.g., soft) or taste (e.g., sweet). Retrieval from semantic memory involving flavour knowledge as in the word "sweet" increased specific activation in the left orbitofrontal cortex which is known to process semantic comparisons among edible items (Goldberg, Perfetti, & Schneider, 2006a).

A number of studies supported a simulation account of memory with retrieval dependent on perceptions by showing temporal overlaps between encoding and retrieval (Kent & Lamberts, 2008). There is not a strict temporal regularity between retrieval and encoding as far as the

ERP evidence shows (Allan, Robb, & Rugg, 2000). However, better memory performance was found in serial recall when retrieval direction (forward vs. backward) matched with the order in which the words were encoded in the first place (J. G. Thomas, Milner, & Hanerlandt, 2003). More direct evidence for temporal similarity between encoding and retrieval comes from Kent and Lamberts (2006). Participants were instructed to retrieve different dimensions of faces such as eye colour, nose shape, mouth expressions etc. Results revealed that features that were quickly perceived were also quickly retrieved.

In addition to the findings from the abovementioned research areas, the historical phenomena of *state-dependent memory* and *context-dependent memory* show that memory retrieval is simulation of the original event. An overlap between the internal state (e.g., mood, state of consciousness) or external context of the individual during encoding and retrieval leads to higher retrieval efficiency (S. M. Smith & Vela, 2001; Ucros, 1989). In one such study, Dijkstra, Kaschak and Zwaan (2007) documented faster retrieval when body positions and actions during retrieval of autobiographical events were similar to the body positions and actions in the original events compared to when body positions and actions were non-congruent. For example, participants were faster to remember how old they were at a concert, if they were instructed to sit up straight in the chair and clap their hands several times during the retrieval. In another intriguing study (Casasanto & Dijkstra, 2010), participants were instructed to tell their autobiographical memories with either positive or negative valence, while moving marbles either upward or downward, which was an apparently meaningless action. However, retrieval was faster when the direction of movement was congruent with the valence of the emotional memory in a metaphorical way (i.e., upward for positive and downward for negative memories).

Lastly, eye movements provide plentiful evidence that retrieval is perceptual recreation of encoding (D. C. Richardson & Spivey, 2000; Spivey & Geng, 2000) and further, these simulations usually predict the success of the retrieval (Johansson & Johansson, 2014; Scholz et al., 2018, 2016). Eye movements in memory simulations are further elaborated in Chapter 2.2.3.

2.1.3 Simulations in language

Language is one of the most influential domains in showing the centrality of simulations in human cognition. The claim of simulation view of language is simple: "Meaning centrally involves the activation of perceptual, motor, social, and affective knowledge that characterizes the content of utterances" (Bergen, 2007, pp. 277-278). Thus, a simulation mechanism is essential to comprehend and remember language.

Switch-cost effects are a clear demonstration of perceptual and affective (re)activation in language. In this paradigm, participants are asked to verify whether a property (e.g., "blender") corresponds to a particular target modality (e.g., "loud" in the auditory modality). The effect is that participants are slower to verify a property in one perceptual modality (e.g., "blender" can be loud - auditory modality) after verifying a property in a different modality (e.g., "cranberries" can be tart - gustatory modality) than after verifying a property in the same modality (e.g., "leaves" can rustle - auditory modality) (Pecher, Zeelenberg, & Barsalou, 2003). A switch-cost occurs between properties with positive and negative valence (e.g., "couple" can be happy, and "orphan" can be hopeless) (Vermeulen, Niedenthal, & Luminet, 2007) and at the sentence level (e.g., "A cellar is dark" in visual modality - "A mitten is soft" in tactile modality) (Hald, Marshall, Janssen, & Garnham, 2011).

Similar switching costs occur when participants switch between actual modalities in perceptual tasks (Masson, 2015). Thus, findings reviewed above support the claim that language is rooted in perceptions and language comprehension can activate these perceptions. Importantly, the same priming effect was not elicited when participants verified semantically associated properties (e.g., "sheet" can be spotless, and "air" can be clean) as opposed to unassociated properties (e.g., “sheet” can be spotless, and “meal” can be cheap) (Pecher et al., 2003). This finding rules out the alternative, computational hypothesis that properties across all modalities are stored together in a single, amodal system of knowledge. Rather, they support perceptual roots of language processing and language-based simulations.

Mental simulations and situation models

Simulations triggered with language are slightly different than the sensorimotor simulations that have been covered so far. Sensorimotor simulations in mental imagery and memory rely on actual sensorimotor experiences (e.g., playing a piano or perceptually encoding an episode). They take place in an offline manner, that is, when the agent “needs” to access perceptual/conceptual information in the absence of original stimulus. Whereas, language-based simulations are activated upon perceiving linguistic stimuli in an online manner. The subject (re)creates perceptual, motor, affective, introspective and bodily states not by actually experiencing them but through linguistic descriptions. Further, language can give rise to simulations of several abstract conceptualisations that go beyond these states. This type of simulation is usually referred to as a *mental simulation* (Zwaan, 1999). Mental simulations can extend into and affect subsequent perceptual/conceptual processing and memory retrieval (discussed below).

It is reasonable to assume that online mental simulation evoked by language and offline simulation in memory and mental imagery share some common architecture. After all, both types of simulations originate from perceptual, motor, affective, introspective and bodily states. That said, the substantial difference between offline and online simulation is conscious effort. Mental simulations based on language are assumed to be inherently involved in language comprehension and thus, triggered automatically and unconsciously (Zwaan & Pecher, 2012). Whereas, offline sensorimotor simulation in memory and mental imagery is often a consequence of effortful, resource-consuming and conscious processes as memory and mental imagery themselves. In line, there is little to no evidence that mental simulation is correlated with the strength of mental imagery (Zwaan & Pecher, 2012).

The idea of mental simulation via language stems from the discovery of *mirror neurons* (Caggiano et al., 1996; Gallese et al., 1996). Mirror neurons are activated in motor regions of the brain by merely observing others executing motor actions (Hari et al., 1998). In a similar fashion, neural correlates were found between the content of what is being read and activated areas in the brain (see Hauk & Tschentscher, 2013; Binkofski, 2010; Pulvermüller, 2005 for exhaustive reviews and Jirak, Menz, Buccino, Borghi for a meta-analysis). In a pioneering study (Hauk, Johnsrude, & Pulvermüller, 2004), participants saw action words referring to face, arm and leg (e.g., lick, pick and kick) in a passive reading task and then, moved their corresponding extremities (i.e., left or right foot, left or right index finger, or tongue). Results showed that reading action verbs activates somatotopic brain regions (i.e., regions corresponding to specific parts of the body) that are involved in the actual movements (see also Buccino et al., 2005). For example, reading the word "kick" or "pick" invokes activation in the specific regions of motor and premotor cortex that control the execution of leg and arm movements respectively. Critically, several fMRI (functional magnetic resonance imaging) studies showed that not only concrete words but also idiomatic expressions involving action words (e.g., "John *grasped* the idea or Pablo *kicked* the habit) (see Yang & Shu, 2016 for a review) and counterfactual statements (e.g., "if Mary had cleaned the room, she would have moved the sofa") (de Vega et al., 2014) elicit similar somatotopic activation in brain. In addition to action words, words in different perceptual modalities activate brain regions associated with the concerned modalities as well. For example, reading odour-related words such as "cinnamon", "garlic" or "jasmine" triggers activations in primary olfactory cortex, the brain region involved in the sensation of smells (González et al., 2006).

Language-based simulations go beyond recreation of perceptual and motor experiences. It is well-documented that reading narratives can form *situation models* (*mental models*) in the minds of the readers (e.g., Speer, Reynolds, Swallow, & Zacks, 2009). Situation models are integrated, situational mental representations of characters, objects and events that are described in narrative (Johnson-Laird, 1983; Kintsch & van Dijk, 1978). They allow readers to imagine themselves in the story by taking the perspective of the protagonist (e.g., Avraamides, 2003). Consequently, situation models give rise to simulations of perceptual, motor and affective states and also abstract structures such as time, speed, space, goals and causations (Speed & Vigliocco, 2016; Zwaan, 1999; Zwaan & Radvansky, 1998).

For instance, Zwaan, Stanfield and Yaxley (2002) evidenced that language comprehenders simulate what the objects described by language look like. In their study, participants read sentences describing an animal or an object in a certain location (e.g., egg in a carton vs. egg in a pan). Thus, the shape of the objects changed as a function of their location, but it is only implied by sentences (e.g., "The egg is in the carton." - whole egg). Even though, a line drawing of the object matching with the shape implied in the previous sentence (e.g., a drawing of a whole egg) improved participants' performance in retrieval of the sentences. Similar results were demonstrated for sentences that imply orientation (e.g., vertical - horizontal) (D. C. Richardson, Spivey, Barsalou, & McRae, 2003), rotation (Wassenburg & Zwaan, 2010), size (de Koning, Wassenburg, Bos, & Van der Schoot, 2017), colour (Zwaan & Pecher, 2012), visibility (Yaxley & Zwaan, 2007), distance (Vukovic & Williams, 2014) and number (Patson, George, & Warren, 2014).

Language can activate simulations of more abstract structures in the same manner. Simulation of time, in particular, is well-documented. For instance, longer chronological distance between two consecutively narrated story events denoted with "an hour later" as compared to "a moment later" leads to longer reading times (Zwaan, 1996). Reading times measured with eye movements were also shown to be longer when reading "slow" verbs (e.g., amble) than “fast” verbs (e.g., dash) (Speed & Vigliocco, 2014). Similarly, Coll-Florit and Gennari (2011) found that judging the sensicality of sentences describing durative states (e.g. “to admire a famous writer”) took longer than non-durative states (e.g. “to run into a famous writer). Several other abstractions can be mentally simulated in the reader’s mind. In one experiment, participants can access the concept of “cake” more easily when they previously read a sentence in which a cake is actually present (“Mary baked cookies and cake”) than when it is not (“Mary baked cookies but no cake”) (MacDonald & Just, 1989). In another experiment, participants simulated the protagonist’s thoughts and they remembered and forgot what the character in the story remembered and forgot (Gunraj, Upadhyay, Houghton, Westerman, & Klin, 2017). In Scherer, Banse, Wallbott and Goldbeck (1991), participants simulated the intended emotions that were cued in characters’ voices.

Mental simulations via language, and situation models play important roles in numerous cognitive tasks transcending language comprehension. Most importantly, simulations are involved in memory for language. Johansson, Oren and Holmqvist (2018) reported that eye movements on a blank screen when participants were remembering a narrative reflected the layout of the scenes described in the text rather than the layout of the text itself. Zwaan and Radvansky (1998) assumed that successful retrieval of what is comprehended would necessarily involve the retrieval of simulations. In accordance with this assumption, there is evidence that the ability to restructure situation models have beneficial effects on memory performance (Garnham, 1981; Magliano, Radvansky, & Copeland, 2012).

Simulation of space with language

Space has a privileged status in human cognition. Coslett (1999) argues that the representation of space in the mind has a fundamental evolutionary advantage because information about the location of objects in the environment is essential for sustenance and avoiding danger. A large body of evidence indicates that young children show sensitivity to spatial concepts and properties starting from the infancy (e.g., Aguiar & Baillargeon, 2002; Casasola, 2008; Frick & Möhring, 2013; Hespos & Rochat, 1997; McKenzie, Slater, Tremellen, & McAlpin, 1993; Örnkloo & Von Hofsten, 2007; Wishart & Bower, 1982). There is also evidence suggesting that development of spatial cognition forms the foundation for subsequent cognitive structures such as mathematical aptitude (Lauer & Lourenco, 2016), creativity (Kell, Lubinski, Benbow, & Steiger, 2013) and notably, language (Levinson, 1992; Piaget & Inhelder, 1969). As a result, there is good reason to assume that language and space are inherently interconnected through the course of cognitive development (e.g., Casasola, 2005; Haun, Rapold, Janzen, & Levinson, 2011; Hespos & Spelke, 2004).

People use language when describing space and spatial language schematises space by selecting certain aspects of a scene while ignoring other aspects (Talmy, 1983). For instance, "across" conveys the information that the thing doing the crossing is smaller than the thing that is being crossed (Tversky & Lee, 1998). However, it does not contain any information about the distance between these things or their shapes. Thereby, language forms spatial representations in mind (H. A. Taylor & Tversky, 1992).

On the other hand, space provides a rich canvas for representing abstraction. Many abstract conceptualisations such as time (Boroditsky & Ramscar, 2002), valence (Meyer & Robinson, 2004), power (Zanolie et al., 2012), numerical magnitude (Dehaene et al., 1993), happiness (Damjanovic & Santiago, 2016), divinity (Chasteen, Burdzy, & Pratt, 2010), health (Leitan, Williams, & Murray, 2015) and self-esteem (J. E. T. Taylor, Lam, Chasteen, & Pratt, 2015) are understood with space (e.g., "powerful is up", "more is up", "happy is up" etc.). Further, space constraints the use of language with gestures and in sign language (Emmorey, 2001; Emmorey, Tversky, & Taylor, 2000). In support of this, both brain imaging (Carpenter, Just, Keller, Eddy, & Thulborn, 1999) and behavioural (Hayward & Tarr, 1995) evidence indicate that there are similarities between spatial and linguistic representations.

Given the central position of space in human mind as briefly discussed above and the intrinsic links between language and space, spatial simulations in language deserve particular attention. Reading narratives can activate simulations of spatial descriptions in a text through situation models. For instance, objects that are described close to a protagonist in a narrative are accessed faster than the objects described as more distant (Glenberg, Meyer, & Lindem, 1987; Morrow, Greenspan, & Bower, 1987). In seminal work, Franklin and Tversky (1990) showed that situation models of space derived from text are similar to the representations of spatial experiences in the real-world and notably, have bodily constraints. Participants in the study read descriptions of scenes and objects in them. Then, they were asked to remember and locate certain objects in a three-dimensional environment. Results showed that objects on the vertical (i.e., head-feet) axis were retrieved faster than objects on the horizontal (i.e., left-right) and sagittal (i.e., front-back) axes. The findings indicate that space in language is simulated with an *ego-centric* perspective rather than an *allocentric* (i.e., object-centred) or a mental transformation perspective. If the participants took an allocentric perspective as in inspecting a picture (in which the subject is not immersed into the environment), all directions would have been equally accessible. On the other hand, if they mentally transformed the described environments, response times would have varied as a function of the mental movement needed to inspect each location. Accordingly, response times would have been shortest for the objects in front of the subject and the accessibility would have decreased in line with the angular disparity from the front. Objects behind the subject, for example, would have been the most difficult to access. Bias for the objects on the vertical dimension suggests that simulation of space with language is body-based. As Franklin and Tversky (1990, p. 64) discuss, the dominant position of a person interacting with the environment is upright due to a number of reasons: First, the perceptual world of the observer can be described by one vertical and two horizontal dimensions (i.e., left/right and front/back). Second, vertical dimension is correlated with gravity, which in an important asymmetric factor in perceiving spatial relations. Thus, vertical spatial relations generally remain constant with respect to the observer. Third, the ground and the sky present stationary reference points on the vertical axis. On the other hand, horizontal spatial relations change frequently. Thus, horizontal dimension depends on more arbitrary reference points, such as the prominent dimensions of the observer's own body. In another experiment using a similar methodology (Avraamides, 2003), it was demonstrated that simulated ego-centric positions are not static but can be automatically updated whenever the reader/protagonist moves in the text, suggesting the motor basis of language. In a recognition memory task, Levine and Klin (2001) showed that a story character's current location was more active in the reader's memory than her/his previous location (see Gunraj et al., 2017). Further, such spatial simulations remained highly accessible even several sentences after last mention, indicating the robustness of these spatial simulations.

There are stable representational mappings between language and space at the sentence level as well. Richardson, Spivey, Edelman and Naples (2001) asked participants to read sentences involving concrete and abstract action verbs (e.g., lifted, offended). They were then asked to associate diagrams illustrating motions on the horizontal (left and right) and the vertical axis (up and down) with the sentences depicting motion events. Substantial agreement was found between participants in their preferences of diagrams for both concrete and abstract verbs within action sentences. For example, participants tended to attach a horizontal image schema to "push", and a vertical image schema to "respect". In a later study, it was evidenced that spatial simulation triggered by a verb affects other forms of spatial processing along the same axis both in a visual discrimination and a picture memory task (D. C. Richardson et al., 2003). Spatial simulations interfered with visual discrimination on the congruent axis and deteriorated performance; however, memory performance was facilitated when the picture to be remembered and the simulated orientation matched (see "Effects of mental simulation" below). The effect was shown for both concrete and abstract verbs. Not only orientation, but upward and downward motion on the vertical axis are simulated via language. In one study (Bergen, Lindsay, Matlock, & Narayanan, 2007), subject nouns and main verbs related with up and down locations interfered with visual processing in the same location. However, the effect was shown in literal sentences implying real space (e.g., “The ceiling cracked” – downward movement for the subject noun, “The mule climbed” – upward movements for the main verb) but not in sentences implying metaphorical space (e.g., “The prices rose”). Bergen et al. (2007) argue that the comprehension of the sentence as a whole, and not simply lexical associations, yield spatial simulations.

However, there is evidence that single words can also trigger simulation of space. Several abstract nouns such as “tyrant” (up) and “slave” (down) invoke simulations of metaphorical spatial locations (e.g., Giessner & Schubert, 2007). There are numerous common nouns in language such as “bird” (up) and “worm” (down) which are associated with actual spatial locations (i.e., *spatial iconicity*). Words denoting spatial locations simulate perceptions of these locations in space. In Zwaan and Yaxley (2003), participants were presented word pairs with spatial associations (e.g., “attic” - “basement”) and asked to decide whether the words are semantically related. Results showed that word pairs in a reverse-iconic condition (i.e., “basement” above “attic”) were judged slower than word pairs in an iconic condition (i.e., “attic” above “basement”). In a similar fashion, it was shown that reading words that occur higher or lower positions in the visual field (e.g., head and foot) hinders the identification of visual targets at the top or bottom of the display (Estes, Verges, & Barsalou, 2008).

Effects of mental simulations

Simulation-based language understanding leads to two main effects on simultaneous or subsequent visual/conceptual processing: compatibility and interference (see Fischer & Zwaan, 2008 for a review). The underlying idea is that if understanding an utterance involves the activation of perceptual, affective and motor representations; then perceptions, emotions and actions that are congruent with the content of the utterances should facilitate visual/conceptual processing and vice versa (Bergen, 2007).

For example, the *action-sentence compatibility effect* demonstrates compatibility/interference resulting from motor simulations in language. In the study introducing the effect for the first time (Glenberg & Kaschak, 2002), participants were presented sensible and non-sensible sentences (e.g., “Boil the air”) and were asked to judge whether the sentences made sense or not. Sensible sentences implied actions either toward the body (e.g., “Open the drawer”) or away from the body (e.g., “Close the drawer”). Response button for identifying the sentence as sensible (i.e., yes button) was either near or far from the participants’ bodies. Results showed that when the implied direction of the sentence and the actual action to press the button matched, participants were faster to judge the sensibility of the sentences. For example, the sentence, “Open the drawer” was processed faster when participants reached the yes button near them, an action that is comparable to opening a drawer. The effect was found not only for imperatives but also for descriptive sentences (“Andy delivered the pizza to you” - toward sentence / “You delivered the pizza to Andy” - away sentence). Notably, sentences describing abstract transfers (“Liz told you the story” - toward sentence / “You told Liz the story” - away sentence) elicited an action-sentence compatibility effect as well. An action-sentence compatibility effect extends to sign language, suggesting that the motor system is involved in the comprehension of a visual-manual language as well (Secora & Emmorey, 2015). Notably, the congruency effect was found relative to the verb’s semantics (e.g., “You *throw* a ball” - away) not relative to the actual motion executed by the signer and perceived by the participant (e.g., “You *throw* a ball” - toward). Along with that, there are meta-reviews and experimental evidence arguing that an action-sentence compatibility effect is generally weak (Papesh, 2015; but see Zwaan, van der Stoep, Guadalupe, & Bouwmeester, 2012) or highly task-dependent (Borreggine & Kaschak, 2006). In sum, the current status of the literature suggests that the factors modulating an action-sentence compatibility effect and in general, effects of language-based simulations should be further specified.

Simulations can also interfere with language comprehension which results in a “mismatch advantage”. For example, Kaschak et al. (2005) demonstrated that participants judge the feasibility of motion sentences (e.g., “The horse ran away from you”) faster when they simultaneously view visual displays depicting motion in the opposite direction as the action described in the sentence (e.g., a spiral moving towards the centre). They concluded that visual processing and action simulation during language comprehension engage the same neural circuits; which, in turn leads to a mismatch advantage. Connell (2007) evidenced a mismatch advantage in the simulation of colour with language. Participants read sentences involving an object which can occur in different colours (e.g., meat can be red when raw and brown when cooked). They were then presented pictures of objects and they had to decide whether the pictured object had appeared in the preceding sentence. Colour of the objects sometimes matched with the descriptions in the sentences (e.g., “John looked at a steak in the butcher’s window” - red steak) and sometimes did not match (e.g., “John looked at a steak in the butcher’s window” - brown steak). Responses were faster when the colour of the object mismatched with the colour implied by the previous sentence.

Why do some studies show a congruency advantage and others an incongruency advantage? This is an important question within the context of the present thesis (see Chapter 6). Kaschak et al. (2005) argue that there are two factors determining match or mismatch advantage in language-based simulations: (1) Temporal distance between the perceptual stimulus and the verbal stimulus to be processed. (2) The extent to which the perceptual stimulus can be integrated into the simulation activated by the content of the sentence. In support of the temporal distance assumption, Borreggine and Kaschak (2006) found that action-sentence compatibility effect arises only when individuals have enough time to plan their motor response as they process the sentence. According to Kaschak et al. (2005), if the verbal information must be processed simultaneously with the perceptual information, a congruency or incongruency advantage may occur, depending on whether linguistic information and perceptual stimulus can be integrated. To be more specific, a congruency advantage is expected if the linguistic and visual stimulus are comparable such as reading the sentence “The egg is in the carton” and seeing a line drawing of a whole egg (Zwaan et al., 2002). However, different perceptual and linguistic stimulus such as reading the sentence “The horse ran away from you” and seeing a spiral moving towards the centre or away from it (Kaschak et al., 2005) result in an incongruency advantage (see also Meteyard, Zokaei, Bahrami, & Vigliocco, 2008).

2.2 I Look, Therefore I Remember: Eye Movements and Memory

2.2.1 Eye movements and eye tracking

Eyes do not flow in a smooth fashion when engaged in visual tasks (Huey, 1908). If you were able to see your gaze on the page or on the digital screen right now, you would notice that your eyes shift from one word to the next as you are reading this sentence. Known as *saccades*, these "jumps" are rapid, short and repeated ballistic (i.e. jerk-like) movements which occur approximately three to four times every second. Saccades abruptly change the point of *fixations*, the periods of eye immobility in which visual or semantic information is acquired and processed (Purves, Augustine, & Fitzpatrick, 2001; D. C. Richardson & Spivey, 2004). In simple terms, individuals internalise the visual world during fixations that are executed between saccades (Bridgeman, Van der Heijden, & Velichkovsky, 1994; Simons & Rensink, 2005). Eye movements are fundamental to visual perception because visual system cannot process the huge amount of available information in the visual world at once. Thus, execution of eye movements allows us to see the world as a seamless whole, although we can only see one region at a time (Buswell, 1936; Yarbus, 1967) due to anatomical limitations (i.e., the total visual field that the human eye covers) and also, limited processing resources (Levi, Klein, & Aitsebaomo, 1985; D. C. Richardson, Dale, & Spivey, 2007).

Fixations have two elemental measures: location and duration. Both measures are highly informative of ongoing cognitive operations. We can see a stimulus clearly only when it falls into the most sensitive area of the retina (i.e., *fovea*) (~2^{o} or 3 to 6 letter spaces), which is specialised for high acuity visual perception (Mast & Kosslyn, 2002; Yarbus, 1967). Thus, eye position (i.e., fixation location) gives valuable information about the location of the attentional "spotlight" (Posner, Snyder, & Davidson, 1980). In other words, fixation location corresponds to the spatial locus of cognitive processing. On the other hand, fixation duration corresponds to the duration of cognitive processing of the material located at fixation (Irwin, 2004, p. 2). Longer fixations suggest higher cognitive load or higher attentional processing demands required by a material or task (Irwin, 2004). The underlying idea behind the link between cognition and fixation is known as *eye-mind assumption* (Just & Carpenter, 1980), which simply posits that the "direction of our eyes indicates the content of the mind" (Underwood & Everatt, 1992). Based on the location and duration of fixations, cognitive processes can be measured and evaluated objectively and precisely during the occurrence of the process in question. There is now a universal consensus on the value of eye movements and eye tracking as a methodology in the investigation of the human mind (e.g., Hyona, Radach, & Deubel, 2003; Just & Carpenter, 1980; Rayner, 1998; Rayner, Pollatsek, Ashby, & Clifton, 2012; Reichle, Pollatsek, Fisher, & Rayner, 1998; Theeuwes, Belopolsky, & Olivers, 2009; Van der Stigchel et al., 2006). Eye tracking methodology provides detailed measures with regard to the temporal order of fixations and saccades, gaze direction, pupil size and time spent on pre-defined regions of the scene. Fixation duration in a certain location relative to other locations is used as the main measure of looking behaviour in the present thesis.

Eye movements can be monitored in various different ways. A *pupil corneal reflection* technique, that is based on high-speed cameras and near infrared light, is the most advanced remote and non-intrusive eye tracking method as of today. An illuminator shines dispersed infrared light to one eye or both eyes. A high-speed video camera captures the infrared reflections coming from the pupil and cornea (i.e., the outer layer of the eye) and transforms them into high-resolution images and patterns pertaining to the position of the eye(s) at any given millisecond. Such an infrared eye tracker can record eye movements quite precisely. Precision offered by an eye tracker is indicated by temporal resolution (i.e., sampling rate) and spatial resolution. Sampling rate shows the frequency of which a tracker samples and determines the position of the eye at a given moment. For example, the eye tracker used in the present thesis (i.e., SR EyeLink 1000) operates at a sampling rate of 1000 Hz, which means that the position of the eye is measured 1000 times every second. Put differently, it produces one sample of the eye position per one millisecond. Spatial resolution refers to the angular distance between successive samples of eye position. Thus, an eye tracker with a higher spatial resolution can detect even the smallest eye movements in a certain interest area. SR EyeLink 1000 has a spatial resolution of 0.25^{o} - 0.50^{o} which means that it can detect and sample eye movements within an angular distance of 0.25^{o} - 0.50^{o}.

There generally exists a spatial difference between the calculated location of a fixation and the actual one. This difference is expressed in degrees of visual angle and reflects the accuracy of eye tracking. If you draw a straight line from the eye to the actual fixation point on the screen and another line to the computed one, the angle between these lines gives the accuracy. Thus, a smaller difference means higher accuracy. Accuracy depends on the screen size and the distance between the participant and the screen. Visual angle is also used to calculate the size of the experimental stimulus as it refers to the perceived size rather than the actual size. These measures of data quality are reported in the methods section of each experiment in accordance with the eye tracking standards and good practices in literature (Blignaut & Wium, 2014; Holmqvist, Nyström, & Mulvey, 2012; D. C. Richardson & Spivey, 2004).

2.2.2 Investigating memory with eye movements

The role of eye movements in evidently visual tasks and processes such as visual perception (Noton & Stark, 1971), reading (Rayner, 1998), visuospatial memory (Irwin & Zelinsky, 2002), visual search (Rayner, 2009) and visuospatial attention (Van der Stigchel et al., 2006) has been widely investigated for many decades and is very well-documented. Eye movements have recently emerged as an alternative means in memory research complementing behavioural measures based on end-state measures (e.g., hit rate, hit latency etc.) (Lockhart, 2000) and brain-imaging studies (Fiser et al., 2016; Gabrieli, 1998; Rugg & Yonelinas, 2003).

It has been known for a long time that previous experience and knowledge of the observer can govern eye movements in addition to the physical properties of the scene and stimulus. For example, many early studies have reported that human observers tend to look at areas of a picture which are relatively more informative to them. Importantly, informativeness rating of a region is modulated by the previous knowledge of the participants in the long-term memory (Antes, 1974; Kaufman & Richards, 1969; Mackworth & Morandi, 1967; Parker, 1978; Zusne & Michels, 1964). Similarly, Althoff and Cohen (1999) reported that previous exposure to a face changes the viewing behaviour and thus, eye movements. In their study, different patterns of eye movements emerged when participants viewed famous versus non-famous faces driven by recognition, fame rating and emotion labelling tasks. Participants made fewer fixations and fixation durations were shorter when viewing famous faces (now known as a *repetition effect*), which suggests lower cognitive load in processing previously experienced stimuli that can be retrieved from memory. Ryan, Althoff, Whitlow and Cohen (2000) took a similar approach: Participants viewed a set of real word images under three conditions: novel (i.e., seen once during the experiment), repeated (i.e., seen once in each block of the experiment) or manipulated (i.e., seen once in original form in the first two blocks and then seen in a slightly changed form in the final block). Participants made fewer fixations and sampled fewer regions when viewing repeated and manipulated scenes compared to novel scenes (i.e., repetition effect). Repetition effect speaks to the link between stability of mental representation and memory-guided eye movements. To illustrate, in Heisz and Shore (2008), the number of fixations gradually decreased with the number of exposures to the unfamiliar faces during a task. There was also evidence for another memory driven eye movement behaviour known as a *relational manipulation effect*: a higher proportion of total fixation time was dedicated to the manipulated regions in the scenes compared with repeated or novel scenes. Further, participants made more transitions into and out of the changed regions of the manipulated scenes than in unchanged (matched) regions of the repeated scenes. Similar paradigms based on eye movements were also used to study memory in non-human primates (Sobotka, Nowicka, & Ringo, 1997), infants (Richmond, Zhao, & Burns, 2015; Richmond & Nelson, 2009) and special populations. For example, Ryan et al. (2000, Experiment 4) did not observe any difference in looking patterns between amnesic patients with severe memory deficits and a control group when both were viewing the repeated images. However, amnesic patients did not look longer at the altered regions when viewing manipulated images, suggesting that amnesia disrupts *relational memory*, i.e., memory for the relations among the constituent elements of an experience. Likewise, in Niendam, Carter and Ragland (2010), schizophrenic patients failed to detect image manipulation, which was shown with eye movements and even though the memory impairment was not evident in behavioural results.

Studies reviewed above suggest relevance of eye movements in memory and importantly, advantages of eye tracking methodology over behavioural, response-based methodologies. (1) Memory-guided eye movements are mostly obligatory, that is, cannot be controlled. For instance, repetition effect reviewed above occurs regardless of the instruction (i.e., whether participants are told just to study all items for later, are explicitly told to pick out the familiar item, or are told to avoid looking at the familiar item) (Ryan, Hannula, & Cohen, 2007, pp. 522-523). (2) Individuals launch memory-guided eye movements whether exposure comes from short term memory (i.e., within the experiment) or from long term memory (i.e., prior to the experiment). (3) Memory-guided eye movements precedes conscious recall. As stated by Hannula et al. (2010), “eye movements can reveal memory for elements of previous experience without appealing to verbal reports and without requiring conscious recollection” (see Spering & Carrasco, 2015 for a comprehensive review; but see Smith, Hopkins, & Squire, 2006). For instance, repetition effect occurs as early as the very first fixation to the item and thus, prior to the behavioural recognition response (Ryan et al., 2007). Similarly, in Henderson and

Hollingworth (2003), gaze durations were reliably longer for manipulated scenes although participants failed to detect changes explicitly.

To conclude, studies making use of eye movements are highly promising as a methodology. They can provide unique information about memory processes, which complement overt behavioural measures and brain imaging (e.g., Hannula & Ranganath, 2009). In fact, eye movements are so representative of memory that mathematical models are able to predict the task that a person is engaged in (e.g. scene memorisation) from their eye movements using pattern classification (Henderson, Shinkareva, Wang, Luke, & Olejarczyk, 2013). It should also be noted that eye movements in memory are not limited to fixation measures or saccadic trajectories. Variation in pupil size (e.g., pupil dilation) and blinks have been used to probe the ongoing processes during retrieval (Goldinger & Papesh, 2012; Heaver & Hutton, 2011; Kahneman & Beatty, 1966; Mill, O'Connor, & Dobbins, 2016; Otero, Weekes, & Hutton, 2011; Siegle, Ichikawa, & Steinhauer, 2008; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004; Vo et al., 2008). A well-established finding is that the pupil dilates as the retrieval becomes cognitively challenging (Goldinger & Papesh, 2012; Kucewicz et al., 2018; Laeng, Sirois, & Gredeback, 2012).

2.2.3 Eye movements in mental imagery and memory simulations

As discussed in Chapter 2.1.1 and 2.1.2, there is mounting evidence showing the neural and behavioural similarities between memory and mental imagery (Albers et al., 2013; Rebecca Keogh & Pearson, 2011). Concordantly, simulation theories of memory within grounded-embodied cognition highlight the connection between memory and mental imagery in that both processes are simulations in essence. That is, memory retrieval/mental imagery is a neural, perceptual and/or motor reinstatement of perception (Borst & Kosslyn, 2008; Buckner & Wheeler, 2001; De Brigard, 2014; Ganis et al., 2004; Kent & Lamberts, 2008; Mahr & Csibra, 2018; Michaelian, 2016b; Norman & O'Reilly, 2003; Pasternak & Greenlee, 2005; Shanton & Goldman, 2010). Eye movements play a crucial role in the simulation thesis of memory and mental imagery because they can illustrate the behavioural reinstatements between perception/encoding and imagery/retrieval. The essential idea behind this imagery - eye movements - memory network holds that eye movements are stored in memory along with the visual representations of previously inspected images and they are re-enacted during memory and visual imagery (Mast & Kosslyn, 2002).

Long before the idea had been proven empirically, many researchers hinted at a possible similarity in saccades between visual perception and imagery (Hebb, 1968; Hochberg, 1968; Neisser, 1967; Schulman, 1983). Hebb (1968) was probably the first researcher who explicitly argued that "if the mental image is a reinstatement of the perceptual process, it should include the eye movements" (p. 470). Brandt and Stark (1997) provided direct empirical evidence for this argument by showing that people do move their eyes during mental imagery and the *scanpaths* (i.e., the sequential order of fixations and saccades, not only their spatial positions) are not random (see also Noton & Stark, 1971 for more on scanpath theory). Instead, they bear striking similarities with the scanpaths during the perception of the original image (Foulsham & Underwood, 2008; Underwood, Foulsham, & Humphrey, 2009) Correspondence between the eye movements in perception and imagery was so robust that it was observed both for auditory (retelling a story) and visual stimuli (depicting a picture) and even when participants were in complete darkness and thus, without any visual information at all during imagery (Johansson et al., 2006). It seems reasonable to assume that spatiotemporal characteristics of visual perception are similar to the mental imagery as eye movements reflect the mental processes during visual inspection. Memory-guided eye movements are also informative in grounding of abstract concepts such as time. In Martarelli, Mast and Hartmann (2017), participants launched more rightward saccades during encoding, free recall and recognition of future items compared to past items (see also Hartmann, Martarelli, Mast, & Stocker, 2014; Stocker, Hartmann, Martarelli, & Mast, 2015).

A majority of the studies investigating the ocular motility in mental imagery and memory have revolved around the role and functionality of eye movements. Whether these eye movements are merely epiphenomenal (i.e., an involuntary by-product of the imagery process) or play an important role and affect the imagery/retrieval performance is an important issue in that it directly taps into the primary question of nonvisual gaze patterns: Why do people move their eyes when forming mental images in the first place? Early studies (Kosslyn, 1980) discussed a potential advantage in vividness if non-random eye movements are systematically employed during mental imagery; yet, they failed to provide experimental evidence, which led to a premature conclusion: Oculomotor movements during imagery were regarded as a mere reflection of the *visual buffer* (Kosslyn, 1980, 1987). A visual buffer is a hypothetical unit which is responsible for holding visual information for a limited time. Nonvisual eye movements in mental imagery were assumed as an additional mechanism for presenting complex scenes on the visual buffer without overloading its capacity (Brandt & Stark, 1997). Thus, eye movements were viewed as passively mirroring the attentional window over the target image during encoding to provide a solution for the cognitive load problem (Irvin & Gordon, 1998).

There is now increasing evidence that eye movements have a relatively more direct role in mental imagery and memory (Bochynska & Laeng, 2015; Hollingworth & Henderson, 2002; Laeng et al., 2014; Mäntylä & Holm, 2006; Stark & Ellis, 1981; Underwood et al., 2009; Valuch, Becker, & Ansorge, 2013). For example, in Laeng and Teodorescu (2002), participants viewed an irregular checkerboard, similar to the one used by Brandt and Stark (1997) or a coloured picture. Then, they were asked to mentally imagine the visual stimuli as they were looking at a blank screen. Percentages of fixation time on certain interest areas and the order of scanning during perceptual phase (i.e., original image) and imagery phase (i.e., blank screen) were highly correlated. But importantly, the strength of relatedness between scanpaths predicted the vividness of mental imagery.

More recent evidence indicates that what is perceptually simulated in memory retrieval or mental imagery is not the order of eye movements (i.e., scanpaths) but rather, the locations of perception. In a visual memory experiment, Johansson, Holsanova, Dewhurst and Holmqvist (2012) found no literal re-enactment during retrieval although suppression of eye movements hindered retrieval accuracy (cf., Bochynska & Laeng, 2015). By challenging the scanpath theory, they deduced that eye movements during retrieval are functional but not one-to-one reactivation of the oculomotor activity produced during perception/encoding (see also Foulsham & Kingstone, 2012 for similar results). Also, in Laeng and Teodorescu (2002), the participants who were not allowed to free scan during imagery phase (i.e., fixed gaze condition) did worse when they were asked to recall the original pattern, which was calculated by the number of squares corresponded to the location of a black square in the grid. Using a similar paradigm in visuospatial memory, Johansson and Johansson (2014) asked participants to view objects distributed in four quadrants at the encoding phase. Participants then listened statements about the direction of the objects (e.g., “The car was facing left”) and were asked to decide whether the statements are true or false. Results showed that participants who were free to look at a blank screen during retrieval had a superior retrieval performance than participants whose eye movements were constrained to a central fixation point. Further, participants whose eye movements were constrained to the previous locations of the objects were more accurate and faster than participants whose eye movements were constrained to a diagonal location as to the previous location of the concerned object.

Studies reviewed above suggest that the human mind encodes eye movements not as they are but in the form of spatial indices, seemingly invisible spatial pointers in space (D. C. Richardson & Kirkham, 2004; D. C. Richardson & Spivey, 2000). Spatial indices link internal representations to objects in the visual world by tapping into space-time information and in turn, trigger eye movements to blank locations during retrieval to reduce working memory demands (Ballard et al., 1997). Therefore, there is no need for a literal recapitulation of gaze patterns because eye movements function as a scaffolding structure with the network of spatial indices for the generation of a detailed image. In other words, spatial indices in the environment which are internalised via eye movements complete the representations “in the head” resulting in a detailed mental image (Ferreira et al., 2008). In an alternative model, O’Regan and Noë (2001) put forward that seeing is a way of acting and eye movements are visual representations themselves in a nod to ecological psychology (Gibson, 1979).

To sum up, current evidence shows that oculomotor activity during memory and mental imagery is not limited to the reconstruction of the original: it is essential to generate mental images. Further, it seems that the role of eye movements is also beyond an automatic and involuntary distribution of limited cognitive sources between the oculomotor activity and memory to alleviate the mental load. Rather, eye movements might serve as an optional, situational strategy in situations where expanding could make a difference for solving the task (Hayhoe et al., 1998; Laeng et al., 2014; J. T. E. Richardson, 1979). In support of this assumption, many task-oriented vision studies have suggested that “the eyes are positioned at a point that is not the most visually salient but is the best for the spatio-temporal demands of the job that needs to be done” (Hayhoe & Ballard, 2005, p. 189). Furthermore, there is also intriguing evidence that these strategic, opportunistic eye movements in goal-directed behaviour are guided by a dopamine-based reward system (Glimcher, 2003; Hikosaka, Takikawa, & Kawagoe, 2000). Thus, eye movements during imagery and memory can be situational and adaptive according to the task demands. For example, in Laeng, Bloem, D’Ascenzo and Tommasi (2014) eye movements during mental imagery concentrated in the salient, information-rich parts of the original image (e.g., head region of an animal picture in the study). Here, it is important to underline that difficulty of the task seems to be the decisive factor. For instance, memory tasks requiring relatively low cognitive load would not need a detailed mental image of the original scene to be solved and thus, retrieval should be challenging in order to observe any memory advantage (Hollingworth & Henderson, 2002; Laeng et al., 2014).

Chapter 3

Experiment 1

Simulating Space when Remembering Words:

Role of Visuospatial Memory

3.1 Motivation and Aims

Spatial simulation within grounded-embodied cognition and cognitive offloading within the extended cognition were outlined and discussed in Chapter 1 and 2. Based on this theoretical background, this chapter describes an experimental study investigating a memory-based looking behaviour (i.e., looking at nothing) which is representative of spatial simulation and cognitive offloading. In general terms, this study aims to investigate how spatial location is simulated following the visual perception of words to support the retrieval of these words.

3.2 Abstract

People tend to look at uninformative, blank locations in space when retrieving information. This gaze behaviour, known as looking at nothing, is assumed to be driven by the use of spatial indices associated with external information. In the present study, we investigated whether people form spatial indices and look at nothing when retrieving words from memory. Participants were simultaneously presented four nouns. Additionally, word presentation was sometimes followed by a visual cue either co-located (congruent) or not (incongruent) with the probe word. During retrieval, participants looked at the relevant, blank location, where the probe word had appeared previously, more than the other, irrelevant blank locations following a congruent visual cue and when there was no cue between encoding and retrieval (pure looking at nothing condition). Critically, participants with better visuospatial memory looked less at “nothing”, suggesting a dynamic relationship between so-called “external” and “internal” memory. Overall, findings suggest an automatic spatial indexing mechanism and a dynamic looking at nothing behaviour for words.

Highlights

- Participants offloaded memory work onto the environment with eye movements when remembering visually and simultaneously presented single words.
- Worse visuospatial memory led to more reliance on the environment during retrieval.

3.3 Introduction

The human mind can anchor spatially-located information to external spatial locations. This mechanism has been expressed within a visual processing model, where the *location* of an object is separated from the *visual feature*s of it (Marr, 1982). This view, expanded into an exhaustive *spatial indexing model* (Pylyshyn, 1989), assumes that the visual system is able to individuate spatial relations before discerning a visual pattern and immediately index the locations of such patterns. In a similar fashion, *spatial registration hypothesis* (Coslett, 1999) holds that perceived stimuli are coded with respect to their location in space. Location, therefore, is a critical constituent of our interactions with the world (van der Heijden, 1993).

Within the spatial indexing (or spatial registering/encoding) model, spatial indices remain attached to a particular object independent of its movements and visual properties. Critically, spatiotemporal continuity (i.e., persistence of spatial “tags” over time) occurs even when the visual information disappears, as often manifested in mental imagery (e.g., Brandt & Stark, 1997). Spatial indices tied to external visual and verbal information trigger eye movements when a mental representation is reactivated. Thus, when retrieving information from memory, people tend to exploit location-based indices and look at the seemingly uninformative, empty locations where the information originally occurred even if location is irrelevant to the task. This behaviour is known as *looking at nothing* (Spivey & Geng, 2000).

In their pioneering study, Richardson and Spivey (2000) documented the use of spatial information and looking at nothing in verbal memory. Four faces randomly appeared on different quadrants of a two by two grid along with four corresponding spoken facts (e.g., "Shakespeare's first plays were historical dramas; his last was the Tempest"). On the next screen, a statement (e.g., "Shakespeare's first play was the Tempest") probed participants' memory for verbal information. During retrieval, there were significantly more looks in the blank quadrant where the face associated with the probed semantic information had been when compared to other quadrants. Thus, people did not just look at any *nothing* when answering the questions. Rather, they looked at an invisible spatial index, which was previously allocated to the information (Spivey & Geng, 2000).

Looking at nothing may be best thought of as an interface between internal and external worlds. Ferreira, Apel and Henderson (2008) proposed an *integrated memory architecture*, where external cues and internal representations work hand in hand to retrieve information as efficiently as possible (see also Richardson, Altmann, Spivey, & Hoover, 2009). More precisely, the integrated memory account combines visual/auditory and spatial information in the external world with visual, linguistic, spatial and conceptual counterparts in the mental world. When part of an integrated representation (linguistic information) is reactivated, the other parts (spatial information) are retrieved as well. In this regard, looking at nothing is also an example of *spatial simulation* (Barsalou, 1999) in that the spatial position where the information is presented is recreated when the information is needed again. Looking at nothing can also be thought as an example of efficient *cognitive offloading* (Risko & Gilbert, 2016), in which the memory work is offloaded onto the world to minimise internal demands.

In the current study, we addressed the looking at nothing triangle, which is composed of actual looking behaviour, spatial indices and mental representations to answer three questions (1) How automatic is spatial indexing? Do individuals automatically index the locations of short and briefly presented linguistic information (e.g., visually and simultaneously presented single words)? (2) How dynamic is spatial indexing and looking at nothing? Can spatial indices be updated with subsequent visual information and how does it affect looking behaviour? (3) Does everybody look at blank locations, or is looking at nothing modulated by certain cognitive capacities such as visuospatial memory span?

3.3.1 Spatial indexing and looking at nothing: automaticity

Looking at nothing typically occurs under two retrieval conditions as shown in the previous studies: (1) People look at blank locations when remembering spoken linguistic information such as factual sentences (Hoover & Richardson, 2008; D. C. Richardson & Kirkham, 2004; D. C. Richardson & Spivey, 2000; Scholz et al., 2018, 2011, 2016). As illustrated above, spoken linguistic information is explicitly associated with a visual object in this paradigm, which we term as *explicit indexing*. In turn, eyes revisit the previous locations of the object (associated with the information) when retrieving the spoken factual information. (2) Looking at nothing also occurs during retrieval of visually presented non-linguistic information such as single objects (Martarelli & Mast, 2013; Spivey & Geng, 2000), arrangement of multiple objects (Altmann, 2004; Johansson & Johansson, 2014) or visual patterns (Bochynska & Laeng, 2015; Laeng et al., 2014). In this case, locations are encoded along with the visual object(s) or patterns.

In the current study, we adopted a different approach to examine the automaticity of spatial encoding of linguistic information. We showed participants four nouns on a grid simultaneously to study for a brief period of time. Then, an auditorily presented word (which could be either among the studied words or not) probed participants' verbal recognition memory while participants were looking at a blank screen. If participants automatically encode location of the words as assumed in spatial indexing hypothesis (*implicit indexing*), they should display looking at nothing behaviour. In other words, we predict more fixations in the now-blank locations of the probe word during retrieval compared to the other, irrelevant blank locations. However, if explicit indexing is required for looking at nothing as shown in the previous studies, there should be same amount of spontaneous looks in the relevant and irrelevant blank locations.

Word locations are encoded in reading (Fischer, 1999; but Inhoff & Weger, 2005), writing (Le Bigot, Passerault, & Olive, 2009) and complex cognitive tasks such as memory-based decision-making (Jahn & Braatz, 2014; Renkewitz & Jahn, 2012; Scholz, von Helversen, & Rieskamp, 2015). However, to what extent spatial encoding is *automatic* is not clear. An automatic process is fast, efficient, unconscious, unintentional, uncontrolled (i.e., cannot be wilfully inhibited), goal-independent and purely stimulus-driven (i.e., cannot be avoided) (Hasher & Zacks, 1979; Moors & De Houwer, 2006). Based on the criteria of automaticity, there is evidence that spatial encoding is an automatic process (Andrade & Meudell, 1993), an effortful process (Naveh-Benjamin, 1987, 1988) and a combination of both (Ellis, 1991). For instance, Pezdek, Roman and Sobolik (1986) reported that spatial information for objects are more likely to be encoded automatically as compared to words.

If participants look at previous locations of the words they are asked to remember, this could provide evidence for the automaticity of spatial encoding in looking at nothing due to the specifics of the present experimental paradigm (i.e., implicit encoding and brief encoding time) and the nature of looking at nothing (i.e., an unintentional and efficient behaviour).

3.3.2 Spatial indexing and looking at nothing: dynamicity

How stable are the spatial indices in looking at nothing? Can they be updated with subsequent visuospatial information? How does updated spatial indices guide eye movements during memory retrieval? Answers to these questions are critical to understand mechanics of spatial indexing and looking at nothing. Thus, we tested whether congruency or incongruency of visuospatial cues between encoding and retrieval stages affects spatial indexing and looking at nothing.

There are studies examining the temporal stability of spatial indices. For example, in Wantz, Martarelli and Mast (2015), location memory for visual objects faded 24 hours after the initial encoding. In consequence, participants looked at relevant, blank locations immediately after the encoding, 5 minutes and 1 hour after the encoding but not after 24 hours. That said, less is known about the spatial stability of indices. In one study (D. C. Richardson & Kirkham, 2004), looking at nothing was reported when the visual information that was associated with the to-be-retrieved information moved and thus, updated the spatial indices. Participants looked at the previous locations of the updated locations rather than the original locations of the previously encoded information, suggesting a flexible and a dynamic spatial indexing mechanism.

In the current study, a visual cue (i.e., a black dot) that was irrelevant to the words and to the task itself was shown between encoding and retrieval stages. The cue was presented either in the same quadrant of the grid or a diagonal quadrant as to the location of the probe word at the encoding stage. There was also a third condition, in which, the participants did not see a cue at all.

A plethora of studies on *Simon effect* (Simon & Rudell, 1967) indicates that spatial congruency between the stimulus and the response results in faster and more accurate response even when the location is irrelevant to successful performance (see Hommel, 2011 for a review). In line, Vankov (2011) presented evidence for a Simon-like effect in spatial indexing and showed that compatibility of irrelevant spatial information benefits memory retrieval (see also Hommel, 2002; Wühr & Ansorge, 2007; Zhang & Johnson, 2004) Participants saw four objects on a 2 x 2 grid (e.g., a line drawing of a guitar, cat, camel and plane) at the encoding phase. Then, they were presented a word denoting either a new object or one of the studied objects (e.g., guitar) in one of the four locations as to the location of the target object; that is, in the same location, a vertical location (above or below the target object), a horizontal location (left or right of the target object) or a diagonal location. Participants were asked to remember whether the object denoted by the word appeared before. The fastest responses were found when the word cue appeared in the same location as to the target object. Participants were the slowest to respond when the word cue was in the diagonal location as to the target object.

In the light of the abovementioned evidence, we predict that (in)congruency between the spatial code attached to the word and the spatial code attached to the visuospatial cue could modulate looking at nothing behaviour. A congruent cue is predicted to emphasize the original location of the probe word and thus, the spatial indice tagging it. In turn, fixations to the relevant, blank locations should be more frequent in congruent cue condition as to no cue condition. On the other hand, an incongruent cue could update the spatial code attached to the word and disrupt looking to blank locations by shifting participants' attention to a diagonal location. Such a pattern would suggest that spatial indexing and looking at nothing for words are dynamic processes that are sensitive to the systematic manipulation of irrelevant visuospatial information.

3.3.3 Looking at nothing and visuospatial memory

The link between mental representations and looking at nothing is critical. One position within the radical grounded-embodied cognition (Chemero, 2011) is that the world functions as an outside memory without the need for mental representations (O'Regan, 1992). According to this view, the external memory store can be accessed at will through visual perception.

As discussed above, the integrated memory account (Ferreira et al., 2008) represents an opposing position within a relatively "traditional" grounded-embodied approach. Accordingly, "internal memory" (mental representations) and so called "external memory" (i.e., the external world internalised via spatial indices and eye movements) work cooperatively in an efficient and goal-directed manner in looking at nothing. To be more precise, "the opportunistic and efficient mind" (D. C. Richardson et al., 2009) exploits external support whenever it needs to minimise internal memory load. In support of this assumption, there is evidence that short-term memory capacity is a reliable predictor of conscious and intentional use of environment in memory tasks (see Risko & Gilbert, 2016 for a review). In one memory study (Risko & Dunn, 2015), offloading (i.e., writing down to-be-retrieved information) was given as an option to the participants. Results revealed that participants with worse short-term memory wrote down the information rather than relying on the internal memory more frequently than the participants with better short-term memory.

In looking at nothing, there is evidence that reliance on the environment increases/decreases in proportion to internal demands. For example, people tend to exhibit less looking at nothing as they are asked to study and recall the same sentences over and over again, suggesting less reliance on external cues as the task becomes easier through repetition (Scholz et al., 2011). Similarly, Wantz, Martarelli and Mast (2015) showed less looks to blank locations with repeated recall without rehearsal as mental representations stabilise in time.

However, not much is known about how individual differences in internal memory map onto the differences in looking at nothing within the scope of integrated memory account. If the opportunistic and efficient cognitive system uses both internal and external cues to access memory traces (D. C. Richardson et al., 2009) and if external cues are used to relieve internal operations (Risko & Dunn, 2015), people with relatively worse visuospatial memory should rely more on the environment during memory retrieval (and vice versa). A correlation between visuospatial memory capacity and looking at nothing could provide further evidence for the integrated memory system by disproving *the world as an outside memory* argument (O'Regan, 1992) and consequently, radical grounded-embodied cognition.

3.3.4 Role of eye movements in memory retrieval

Another fundamental issue is whether looks occur to blank regions that are associated with information facilitate the retrieval of this information. This issue taps into a seemingly simple question with regard to the very nature of memory-guided eye movements: Why do people look at nothing? Role and functionality of eye movements in memory retrieval have been highly controversial (see Ferreira et al., 2008; Mast & Kosslyn, 2002; Richardson et al., 2009 for discussions). First studies did not present any evidence for improvement in memory with looks to blank spaces (Hoover & Richardson, 2008; D. C. Richardson & Kirkham, 2004; D. C. Richardson & Spivey, 2000; Spivey & Geng, 2000; Vankov, 2011). Initial failure to demonstrate memory enhancement lead to the preliminary conclusion that eye movements only co-assist the retrieval process as a by-product (Spivey, Richardson, & Fitneva, 2004).

There is now growing evidence that gaze position can play a functional role in memory retrieval. For example, Laeng and Teodorescu (2002) reported that participants who viewed an image and looked at the blank screen freely (free perception & free retrieval) were more accurate in answering the retrieval questions those whose gaze were restricted to the central fixation point (free perception & fixed retrieval) (see also Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Laeng et al., 2014 for memory advantage in free gaze compared to fixed gaze). In a similar gaze manipulation paradigm, participants who were instructed to look at relevant, blank regions were more accurate in judging statements about visual objects (Johansson & Johansson, 2014) and verbal information (Scholz et al., 2018, 2016) than the participants who were instructed to look at a diagonal location as to the original location of the object or object associated with verbal information.

The current study was not designed to test the role of looking behaviour in memory. That is, eye gaze at retrieval was not manipulated as in the studies reviewed above. Rather, we analysed the functionality of looking at nothing by using the fixation percentage in the relevant quadrant (i.e., looking at nothing) as a predictor of hit rate and hit latency within mixed-effects models. If looks to the relevant, blank locations predict recognition memory for visually presented single words, it might provide tentative evidence for the facilitatory role of gaze position in memory.

3.4 Method

3.4.1 Participants

The experiment was carried out with forty-eight students at the University of Birmingham (six males; M_{age} = 19.92, *SD* = 1.96, range: 18 - 27). 96% of the participants were psychology students. All participants were monolingual native speakers of British English as determined with the Language History Questionnaire (version 2.0; Li, Zhang, Tsai, & Puls, 2013). Participants reported normal or corrected-to-normal vision, no speech or hearing difficulties and no history of any neurological disorder. They received either £6 (*n* = 12) or course credit (*n* = 36) for participation. All participants were fully informed about the details of the experimental procedure and gave written consent. Post-experiment debriefing revealed that all participants were naïve to the purpose of the experiment. No participant was replaced.

3.4.2 Materials

There were 192 trials involving 864 unique nouns in total. Trials were evenly divided into two groups (*n* = 96) as experimental (positive probe) trials and fillers. Probe words in the experimental trials were among the four study words in the encoding phase, whereas a different, not seen, word was probed in fillers. Words in the experimental trials (*n* = 384) were drawn from the extensions of Paivio, Yuille and Madigan norms for 925 nouns (J. M. Clark & Paivio, 2004). The word pool was filtered to exclude words shorter than 3 letters and longer than 6 letters. Imageability, frequency (the CELEX database; Baayen, Piepenbrock, & Gulikers, 1995; and logarithmic values of occurrences per million in Kučera & Francis, 1967), age of acquisition, concreteness, availability (Keenan & Benjafield, 1994), length in letters and number of syllables were identified as major predictors of verbal memory (Rubin & Friendly, 1986) and used to control the experimental stimuli.

The subset was then grouped into quadruples and trial sets were identified. Words within quadruples were matched on age of acquisition, availability, concreteness, imageability, length in letters, log frequency and number of syllables (all *SD*s < 2.00 and all *SE*s < 1.00). Words were further controlled so that no word started with the same letter, rhymed or related semantically with any other in the quadruple. Monosyllabic, disyllabic and trisyllabic words were evenly distributed [e.g., (3, 3, 3, 3), (1, 2, 1, 2) or (3, 2, 3, 2) etc.]. The word in each trial set with the median imageability value was selected as the probe among four words leaving the others as distractors (see Rubin & Friendly, 1986). Welch's t-tests revealed no significant difference between the probe and distractor words in in frequency, length in letters or number of syllables (all *p*s > .05). Thus, any word among the four words in each trial set was as likely to be remembered as any other word. Words in filler trials were drawn from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). They were also controlled to develop a consistent stimuli set. Words were grouped into quintuples and matched on log frequency in CELEX database (all *SD*s < 0.60 and all *SE*s < 0.30).

Finally, we formed 192 unique mathematical equations [e.g., (2*3) - (2+3) = 1] to present as memory interference between encoding and retrieval phases (see Conway & Engle, 1996 for a similar design). Half of the equations were correct. Incorrect equations were further divided into two equal groups: The results were either plus or minus one of the correct result.

3.4.3 Apparatus

Stimuli were presented on a TFT LCD 22-inch widescreen monitor operating at 60 Hz with a resolution of 1680 x 1050 pixels (501.7 mm x 337.4 mm). The monitor was placed 640 mm in front of the participant. A chin and forehead rest was used to reduce head movements. Participants' eye movements were monitored using SR EyeLink 1000 (sampling rate: 1000 Hz, spatial resolution < 0.5°, http://sr-research.com/eyelink1000.html). Viewing was binocular but only the left eye was monitored. Auditory material was produced by a native female speaker of British English in a sound attenuated room and recorded using Audacity (version 2.1.10, https://www.audacityteam.org). Participants responded (yes/no they had seen the word) by pressing one of two keys on a standard keyboard. Eye movement data were extracted using the SR EyeLink Data Viewer (version 2.4.0.198, https://www.sr-research.com/data-viewer/). No drift or blink correction procedure was applied.

Data were analysed and visualised in R programming language and environment (R Core Team, 2017). Mixed-effects models were constructed with *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). Significance values of the likelihood tests and coefficients in models were computed based on the t-distribution using the Satterthwaite approximation with *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015).

3.4.4 Procedure

Eye tracking started with a standard nine-point calibration and validation, which confirmed high data quality (average calibration error < 1° and maximum calibration error < 1.50°). As spelled out in detail below, each trial was composed of five consecutive phases: (1) fixation (2) encoding, (3) cueing, (4) interference and (5) retrieval (See Figure 3.1). The task was to decide whether an auditorily presented word had appeared before or not (i.e., yes/no verbal recognition memory test). As soon as the participants made yes/no judgement by hitting one of the response buttons, the trial ended, and a new encoding phase began.

(1) Fixation: A fixation cross appeared at the centre of the screen for 500 ms. **(2) Encoding**: Participants were presented four words on a 2 x 2 grid for 1600 ms. Words (Times New Roman, font size = 40) were centrally placed in rectangular boxes (285 x 85 in pixels, 7.6° x 2.4° of visual angle). By using boxes during encoding and retrieval, we aimed to enrich the spatial context in order to evoke more reliance on the space and thereby, observe looking at nothing when remembering short verbal information as words (see Spivey & Geng, 2000 for the effect of spatial context). **(3) Cueing**: A flashing black dot appeared in cue trials for 1000 ms either in the same (congruent cue) or in the diagonal quadrant (incongruent cue) as the original location of the probe word in the encoding phase. There was also a third condition where no cue was presented between encoding and interference. Cue condition was a within-subjects variable and three cue conditions were randomly presented in a session. That said, an equal number of random participants (n = 16) saw the same probe word with a congruent cue, an incongruent cue or without any cue. **(4) Interference**: Participants were exposed to retrospective memory interference which was irrelevant to the main task. We expected to push out old information (i.e., encoded words) from the episodic buffer (Baddeley, 2000) and encourage participants to depend on spatial indices for the retrieval of words without using explicit indexing (see Martarelli, Chiquet, Laeng, & Mast, 2017 for a similar paradigm). Hence, participants were presented a mathematical equation and asked to identify whether the equation was correct or not within 10,000 ms. **(5) Retrieval**: The probe word was auditorily presented as participants looked at the blank grid with empty boxes. There was a 500 ms gap at the between the presentation of the blank retrieval screen and the sound file was played. Participants were asked to make an unspeeded yes/no judgement to determine whether they had seen the probe word among the four words shown in the encoding phase within 10,000 ms (or they timed-out).

The order of trials and equations were fully randomised independent of each other. The location of all words in all conditions was counterbalanced with Latin Square design to control gaze biases so that each word appeared an equal number of times in each location of the grid. The experiment was divided into four equal blocks with 48 trials in each block and there was a short pause between blocks. A typical session lasted approximately 60 minutes, including consent and setting up the eye tracker. Overall accuracy in interference equations and in the recognition memory test for words were 86% and 81% respectively, suggesting that participants attended to the task with high concentration.

Following the experiment, a computerized version of the Corsi block-tapping task (Corsi, 1972) operated on PEBL (Psychology Experiment Building Language, version 0.13, test battery version 0.7, http://pebl.org) (Mueller & Piper, 2014) was used to measure visuospatial short-term memory.

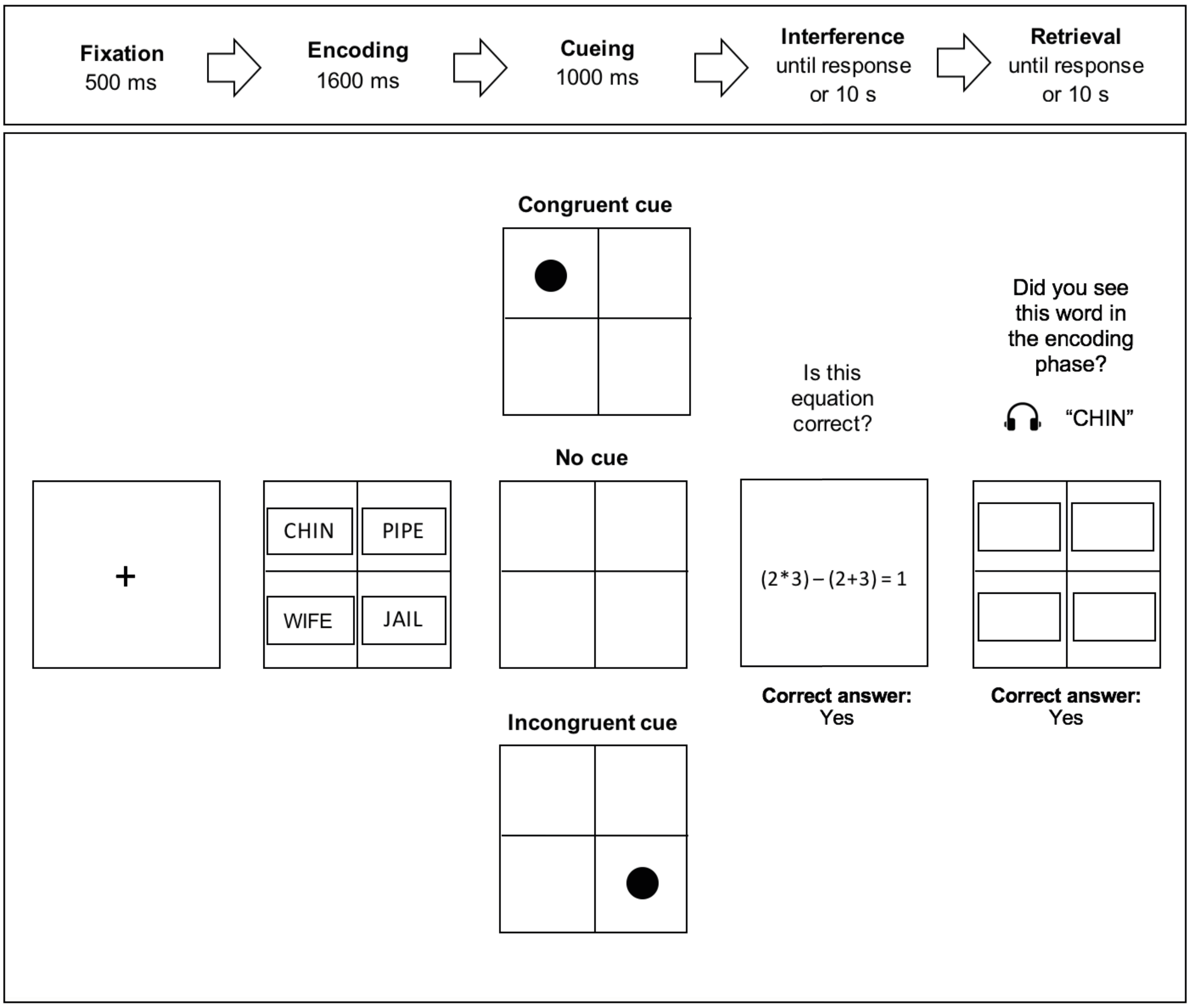

Figure 3.1 A schematic illustration of the temporal order of events in an example trial showing three different cue conditions. In this example, the relevant quadrant is the top left location, where the probe word (i.e., CHIN) appears.

3.5 Results

3.5.1 Measures

Results were analysed in two parts as memory performance and looking behaviour.

Memory performance: Hit rate and hit latency rate were used as measures of memory performance. Hit rate was the proportion of yes trials to which the participants correctly responded yes. Hit latency was the time in milliseconds between the onset of auditory presentation of the probe word and correct keyboard response. Participants were not instructed to make speeded response in the current paradigm. Nevertheless, hit latencies were reported to verify and complement hit rate.

Looking behaviour: Fixation percentage was used as the main gaze measure and dependent variable as in previous looking at nothing studies discussed above (e,g., Wantz et al., 2015). Fixation percentage (or fixation frequency) is the percentage of fixations in a trial falling within a particular interest area in proportion to total fixations in a trial. Thus, it was computed by dividing the number of fixations on each quadrant to the total number of fixations during the retrieval phase (see Wenzel, Golenia, & Blankertz, 2016 for a similar computation and use of fixation frequency).

Words in the study were of varying lengths and thus, had different presentation durations. Fixation percentage was purposefully chosen as it is immune to such differences in durations. Further, we assumed that fixations rather than the time spent on particular region (i.e., dwell time per quadrant) are important for the link between memory and eye movements. Fixation-based measures are reliable indicators of memory load and attention in a given location (e.g., Just & Carpenter, 1980; Meghanathan, van Leeuwen, & Nikolaev, 2015). Hence, we preferred fixation percentage over dwell time percentage as a more refined indicator of looking at nothing[2]. Accordingly, we expected that participants would fixate on the relevant quadrant to derive support from the environment.

Four rectangular interest areas corresponding to the quadrants were identified. All interest areas were of the same size (502 x 368 in pixels, 13.4° x 10.6° of visual angle). They framed the rectangular boxes that words were presented in (see Figure 3.1) and were not contiguous (see Jahn & Braatz, 2014 for a similar arrangement). Interest areas occupied 93.58% of the

[2] The same analyses were performed with dwell time percentages as well and findings were consistent with the analyses based on fixation percentages reported here.

total screen area. A circular interest area with a diameter of 40 pixels (1.1° of visual angle) was defined at the centre of the grid. Participants' head were positioned on a head and chin rest to minimise head movements and we assumed that looking at the centre was the baseline looking behaviour in contrast to the looking at the relevant quadrant. Negative correlation between looks to the centre and the relevant quadrant confirmed this inverse relationship; $r_s\,(46) = -.73$, $p < .0001$.

Proportion of fixations accrued on the interest areas during the retrieval phase (from the onset of auditory presentation of the probe word until the participant's response) were calculated. Fixations were a minimum duration of 40 ms. First fixations and fixations outside the interest areas (7.91%) were omitted. Only hits (i.e., correct responses) in yes trials were included in the fixation analyses. Fixation percentages allocated to the three quadrants that did not contain the target probe word were averaged into one and analysed against the relevant quadrant in which the probe word was seen.

3.5.2 Mixed-effects modelling

Data were analysed using linear and binomial logit mixed-effects modelling. Visual inspections of residual plots did not reveal any obvious deviations from homoscedasticity or linearity. Linear models were fit for continuous target variables (hit latency and fixation percentage). Binomial models were fit for categorical target variables (hit rate) and with *bobyqa* optimiser to prevent non-convergence. Participants and items were treated as random effects to explain by-participant and by-items variation (Baayen, Davidson, & Bates, 2008).

We started fitting models by building the random effects structure and followed a maximal approach. That is, random effects were included as both random intercepts and correlated random slopes (random variations) as long as they converged and were justified by the data (Barr, Levy, Scheepers, & Tily, 2013). Random intercepts and slopes were included even if they did not improve the model fit in order to control for possible dependence due to repeated measures or order effects. In particular, imageability and word length among the lexico-semantic variables were selected to add as random slopes as long as the models converge. Random effects structure was simplified step by step as per the magnitude of the contribution of a random effect to the explanation of the variation in the data. That is, the random effect with the weakest contribution was dropped first and if necessary, the structure was further reduced accordingly.

Contribution of a fixed effect was investigated by comparing a full model containing the effect in question against a reduced model in which only that effect was removed, or a null model without any fixed effects. Compared models had the same random effects structure (Winter, 2013).

3.5.3 Memory performance

Hit rate

We analysed whether there was a difference in hit rate across congruent and incongruent cue conditions. Fixed effect was cue location with two levels (congruent and incongruent cue). Imageability was added as random slopes into participants. Imageability and word length were added as random slopes into items. Cue location did not improve the model fit when compared against a null model; $\chi^2(1) = 0.01$, $p = .91$. In other words, participants retrieved the probe words in incongruent cue condition (mean hit rate = 81%) as accurately as congruent cue condition (mean hit rate = 81%). Cue location did not improve the model fit either when no cue condition (mean hit rate = 82%) was included; $\chi^2(2) = 0.48$, $p = .79$.

Hit latency

Linear mixed-effects models were fit to identify any difference in hit latency between cue conditions. Fixed effect was cue location with two levels (congruent and incongruent cue). Imageability, word length and cue location were added as random slopes into participants. Imageability and word length were added as random slopes into items. As in hit rate, likelihood tests indicated that there was no difference in hit latency between congruent (mean hit latency = 1807.48 ms) and incongruent (mean hit latency = 1830.94 ms) cue condition; $\chi^2(1) = 1.47$, $p = .23$. Results did not change when no cue condition (mean hit latency = 1842.84 ms) was included; $\chi^2(2) = 2.59$, $p = .27$.

Effect of visuospatial memory on memory performance

The effect of visuospatial memory capacity of participants as measured by Corsi-block tapping test on hit rate and hit latency was examined. As reported above, we did not find any differences in hit rate or hit latency across cue conditions. Thus, mixed-effects models including all cue conditions were fit. Fixed effect was Corsi-block tapping score.

Hit rate: Imageability was added as random slopes into participants. Corsi-block tapping score improved the model fit; $\chi^2(1) = 9.39$, $p = .002$. Participants with better visuospatial memory retrieved the probe words from memory more accurately; $\beta = 0.01$, $z = 3.19$, $p = .001$.

Hit latency: Word length were added as random slopes into participants. Imageability and word length were added as random slopes into items. Corsi-block tapping score did not improve the model fit; $\chi^2(1) = 0.55$, $p = .46$.

3.5.4 Looking behaviour

Looking at nothing for words

First, we examined whether there was a difference in spontaneous looking times between relevant and irrelevant quadrants during the retrieval phase. The target variable was fixation percentage in the correctly answered yes trials. Fixed effect was quadrant with two levels (relevant and irrelevant quadrant). Imageability was added as random slopes into participants. Imageability and word length were added as random slopes into items. Likelihood tests showed that quadrant significantly improved the model fit; $\chi^2(1) = 22.85$, $p < .0001$. Overall, participants looked significantly more at the relevant quadrant as compared to the irrelevant quadrant when retrieving probe words from memory; $\beta = 0.03$, $t = 4.78$, $p < .0001$.

Effect of visuospatial interference on looking at nothing

We examined looking at the relevant and the irrelevant quadrants within congruent, incongruent and no cue conditions separately to specify the effect of visuospatial interference on looking at nothing (see Figure 3.2). The target variable was fixation percentage in correctly answered yes trials. Fixed effect was quadrant with two levels (relevant and irrelevant quadrant). Imageability was added as random slopes into participants; imageability and word length were added as random slopes into items in all models.

Congruent cue condition: Quadrant improved the model fit; $\chi^2(1) = 27.51$, $p < .0001$. Participants looked significantly more at the relevant quadrant compared to the irrelevant quadrant in congruent cue condition; $\beta = 0.05$, $t = 5.25$, $p = .0001$.

No cue condition: Quadrant improved the model fit with a smaller magnitude compared to congruent cue condition; $\chi^2(1) = 5.00$, $p = .03$. Participants looked significantly more at the relevant quadrant compared to the irrelevant quadrant in no cue condition; $\beta = 0.02$, $t = 2.24$, $p = .03$.

Incongruent cue condition: Quadrant did not improve the model fit; $\chi^2(1) = 0.61$, $p = .43$. Participants did not look significantly more at the relevant quadrant during the retrieval phase; $\beta = 0.007$, $t = 0.78$, $p = .43$. Models including quadrant with three levels (i.e., relevant and irrelevant quadrant and central interest area) indicated that participants did not look at any region more than the other (ps > .05) in incongruent cue condition.

Difference between conditions: Models with fixation percentage in the relevant quadrant as the target variable were fit by taking the no cue condition as the baseline condition. Results revealed that participants looked at the relevant, blank locations more frequently in congruent cue condition as to no cue condition; $\beta = -0.03$, $t = 2.48$, $p = .01$. There was not such a difference between no cue and incongruent cue ($p = .95$).

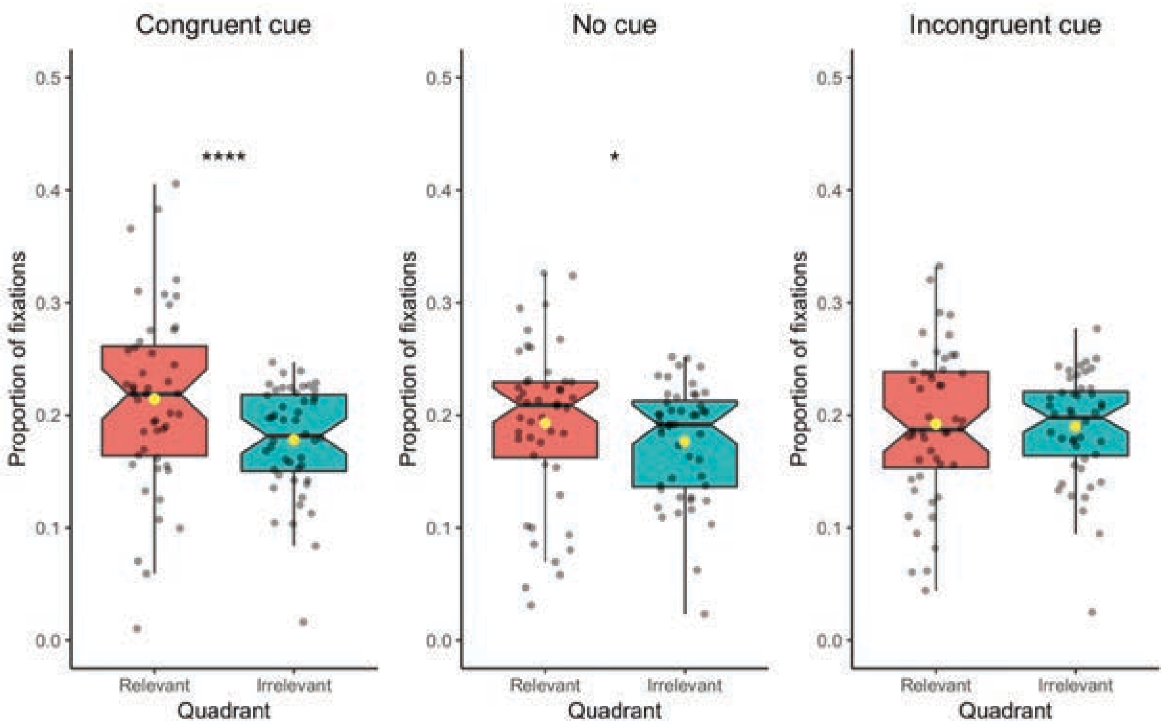

Figure 3.2 Proportion of fixations in the relevant and irrelevant quadrants across three cue conditions. Values on the y axis correspond to percentages (e.g., 0.1 = 10%). Notched box plots show median (horizontal line), mean (yellow dot), 95% confidence interval of the median (notch), interquartile range (the box), the first and the third quartiles (lower and upper ends of the box) and ranges (vertical line). Grey dots represent data points. * $p \leq .05$, **** $p \leq .0001$

Visuospatial memory and looking behaviour

Visuospatial memory as a predictor of looking at nothing

The effect of visuospatial memory capacity of participants as measured by Corsi-block tapping test on looking at the relevant, blank quadrant was investigated. The target variable was fixation percentage in the relevant quadrant in correctly answered yes trials. Fixed effect was Corsi-block tapping score. Word length was added as random slopes into participants; imageability was added as random slopes into items in all models.

Congruent cue condition: Corsi-block tapping score did not predict looks in the relevant quadrant; $\chi^2(1) = 1.07$, $p = .30$ or irrelevant quadrant; $\chi^2(1) = 1.50$, $p = .22$.

No cue condition: Corsi-block tapping score did not predict looks in the relevant quadrant; $\chi^2(1) = 0.30$, $p = .58$ or irrelevant quadrant; $\chi^2(1) = 1.52$, $p = .22$.

Incongruent cue condition: Corsi-block tapping score improved the model fit for fixation percentage in the relevant quadrant; $\chi^2(1) = 4.83$, $p = .03$ but not the irrelevant quadrant; $\chi^2(1) = 0.67$, $p = .41$. Participants with better visuospatial memory looked less at the relevant quadrant during memory retrieval when there was an incongruent cue between encoding and retrieval phases; $\beta = -0.0009$, $t = 2.28$, $p = .03$.

Correlation between visuospatial memory and looking behaviour

We tested the correlation between visuospatial memory as a function of Corsi-block tapping test and fixations to relevant, irrelevant and central interest areas under three different cue conditions (see Figure 3.3).

Relevant quadrant: There was a significant, negative correlation between visuospatial memory capacity and fixation percentage in the relevant quadrant under incongruent cue condition; r_s (46) = -.37, p = .009. Participants with better visuospatial memory tended to look less at the relevant quadrant when there was an incongruent cue between encoding and retrieval phases. There was not such a correlation within congruent; r_s (46) = .20, p = .18 or no cue conditions; r_s (46) = -.18, p = .22.

Irrelevant quadrant: There was a significant, negative correlation between visuospatial memory capacity and fixation percentage in the irrelevant quadrant under the no cue condition; r_s (46) = -.29, p = .05. Participants with better visuospatial memory tended to look less at the irrelevant quadrant when there was not any cue between encoding and retrieval phases. There was not such a correlation within congruent; r_s (46) = -.26, p = .07 or incongruent cue conditions; r_s (46) = -.16, p = .27.

Central interest area: There was a significant, positive correlation between visuospatial memory capacity and fixation percentage in the central interest area under congruent cue condition; r_s (46) = .39, p = .006, no cue condition; r_s (46) = .30, p = .04 and incongruent cue condition; r_s (46) = .33, p = .02. All conditions combined, participants with better visuospatial memory tended to look more at the central interest area; r_s (46) = .30, p = .04.

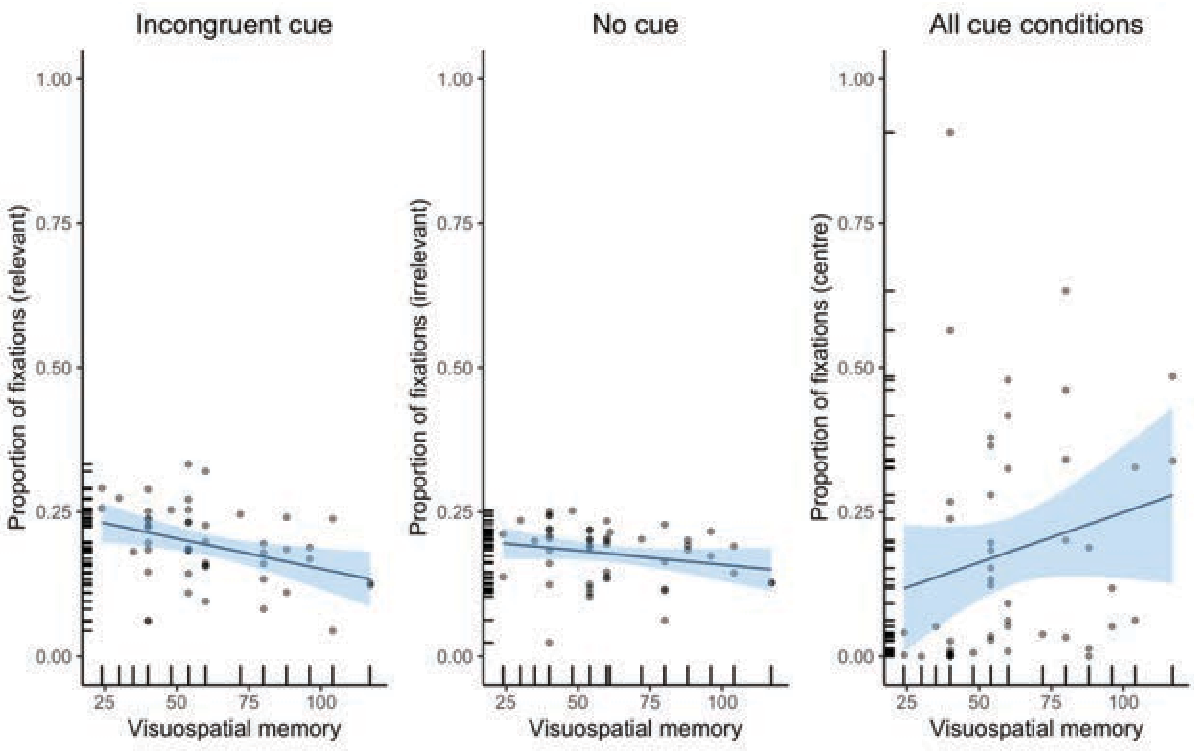

Figure 3.3 Scatterplots showing the correlations between visuospatial memory as a function of Corsi-block tapping score (higher score means better visuospatial memory) and fixation percentage in the relevant, irrelevant and central interest area. Values on the y axis correspond to percentages (e.g., 0.25 = 25%). Scatterplot has a linear regression line. Blue band around the line represents 95% confidence interval. Tassels at the x and y axis illustrate the marginal distribution of data along visuospatial memory and fixation percentage.

Functionality of looking at nothing

The current experiment was not designed to test the functionality of looking behaviour in memory. Nevertheless, we examined whether memory performance (hit rate and hit latency) was predicted by the proportion of fixations in the relevant, blank locations. Imageability was added as random slopes into participants. Imageability and word length were added as random slopes into items. Looks to relevant, blank locations predicted hit rate; $\beta = 0.54$, $z = 2.28$, $p = .02$ in congruent cue but not in no cue; $\beta = 0.31$, $z = 1.22$, $p = .22$. Hit rate also predicted looking at nothing in congruent cue condition only; $\beta = 0.05$, $t = 2.29$, $p = .02$. Looks to relevant, blank locations did not predict hit latency in neither of the cue condition (congruent cue; $\beta =$ -44,74, $t = 0.66$, $p = .51$, no cue; $\beta = 53.56$, $t = 0.70$ $p = .48$). Interaction between Corsi-block tapping score and fixation percentage in the relevant quadrant did not predict hit rate or hit latency (*p*s > .05).

3.6 Discussion

The purpose of the current study is to shed light on the nature of spatial indexing and looking at nothing mechanisms and particularly, to investigate the relationship between internal and external memory within memory for language. To this end, we asked three questions as to automaticity of spatial indexing, dynamicity of spatial indexing and looking at nothing and the effect of individual differences in visuospatial memory on looking behaviour.

3.6.1 Looking at previous word locations

Results showed there were significantly more fixations in the relevant, blank region where the probe word appeared at the encoding stage relative to other, irrelevant blank regions during memory retrieval. In other words, participants looked at nothing when retrieving simultaneously and visually presented single words. Our results in the congruent and no cue conditions were in line with the previous studies evidencing looking at nothing when remembering verbal information (Hoover & Richardson, 2008; D. C. Richardson & Kirkham, 2004; D. C. Richardson & Spivey, 2000; Scholz et al., 2018, 2011, 2016). That is, we replicated the *corresponding area effect* (Wantz et al., 2015). The novelty of this study lies in the linguistic information to be retrieved and how it is encoded and remembered. As discussed in introduction, people saw "fact-teller" objects along with the spoken information in the previous studies which document looking at nothing for language memory (e.g., Richardson & Spivey, 2000). Thus, participants associated verbal information with external visual information. Such an explicit indexing might have motivated participants to rely on the environmental sources. Whereas, linguistic information was not explicitly associated with any visual object in the current study. Further, words appeared in the four cells of the grid at the same time (see Vankov, 2011). Lastly, memory was not probed with details about factual information or correct/incorrect statements but in a simple recognition memory test. In such a relatively minimal and ecologically valid retrieval scenario, participants offloaded memory work onto the environment by simulating locations unintentionally when retrieving linguistic information from memory. It is also important note that looking at nothing in the present study occurred following an intervening task (i.e., judging a maths equation). Simulating spatial locations following a demanding task might suggest that looking at nothing is a not mere residual of the encoding process but rather, an efficient means of memory retrieval (see Renkewitz & Jahn, 2012). To put in a nutshell, our findings suggest that looking at nothing could be a more robust and ubiquitous behaviour than previously documented.

One limitation of the current study could be the use of boxes. We aimed to enrich the spatial context on the screen by following the methodology in Spivey and Geng (2000) by placing words in rectangular boxes at the encoding stage in both experiments. Importantly, participants were asked to remember the probe word while looking at a retrieval screen with boxes without the words in them. This methodology allowed us to identify narrower and thus, more specific interest areas (i.e., boxes) than quadrants of the grid. However, to what extent remembering words while looking at a screen with empty boxes meets the definition of looking at "nothing" in strict terms can be discussed. Replication studies are necessary to ascertain that word locations can be registered, simulated and referred back to via eye movements without any contextual enrichment such as boxes (see Chapter 6 for a methodology where words are not placed in boxes).

3.6.2 Indexing word locations

Visuospatial cues affected spatial indexing and thus, looking at nothing in line with our predictions. Participants looked at relevant, blank locations in congruent cue condition, that is, when the cue appeared in the same location as to the probe word. Importantly, there were also more looks in the relevant quadrant when there was no cue between encoding and retrieval stages (pure looking at nothing). Findings from no cue condition suggest that looking at nothing is not driven by mere attentional shift. Rather, eye movements in the present study resulted from the spatial indices associated with words and thus, were governed by memory for language. On the other hand, looking at nothing did not occur when the visuospatial cue appeared in a diagonal location as to the original location of the probe word (i.e., incongruent cue condition).

Results indicate that participants formed spatial indices corresponding to simultaneously presented single words even though locational information was not required in the memory task. Spatial indices were formed for subsequent cues as well. Emergence and magnitude of looking at nothing were determined by the relationship between the spatial indices for words and cues. Congruent cues reinforced the encoded locations and amplified the corresponding area effect as expected. In turn, participants looked more at the relevant locations in congruent cue condition as compared to no cue condition. In contrast, incongruent cue functioned as interference. When spatial indices associated with words and visual cues did not match, the initial index attached to the word was updated. Consequently, eye movements to the relevant, blank location were disrupted. It is important to note that participants did not look at any blank region (relevant, irrelevant or centre) more than the other regions in incongruent cue condition. Such a behaviour suggests that spatial codes corresponding to words and visuospatial cues were in competition when they did not refer to the same location.

We can conclude that word locations were registered in all cases. Participants were given only 1600 ms to study four words leaving 400 ms for each word. Thus, we can argue that word locations were encoded almost instantaneously upon the presentation. Further, locations were encoded unintentionally suggested by the fact that participants were naïve to the purpose of the experiment and they were not instructed to remember word locations (cf., Andrade & Meudell, 1993; Naveh-Benjamin, 1988). Informal interviews with the participants after the experiment suggested that locations were indexed without awareness. In keeping with this, it appears safe to argue that spatial indexing mechanism in the current study meets most of the automaticity criteria (Moors & De Houwer, 2006). In this regard, our results contrast with Pezdek et al. (1986), which shows automatic spatial encoding for objects but not words.

In conclusion, we present that not only the existence but also the magnitude of looking at nothing is determined by the strength and stability of spatial encoding. Although spatial indexing and looking at nothing are inherently different processes, they are both linked to each other in a dynamic relationship.

3.6.3 Looking at nothing and visuospatial memory

The chief finding of the study is the relation between visuospatial memory capacity and the tendency to look at blank locations. To our knowledge, this is the first direct evidence showing individual differences in looking at nothing. We showed this relation in predictive and correlational analyses. There was a positive correlation between visuospatial memory measured with Corsi-block tapping test and fixation percentages in the central interest area during retrieval in all cue conditions. Higher visuospatial memory predicted less looking at nothing under incongruent cue condition. In line, there were negative correlations between visuospatial memory and fixations in relevant (within incongruent cue condition) and irrelevant locations (within no cue condition).

Taken together, participants with better visuospatial memory, thus richer internal sources, looked more at the centre of the screen rather than looking at relevant (or irrelevant) locations. Central interest area was the initial and thus, default looking position prompted by a central fixation cross shown before each trial. Given that participants' head was stabilised on the chinrest, we assume that participants with better visuospatial memory who "looked" at the centre of the screen; in fact, did not look at any specific area. In other words, they sustained their attention on the internal sources rather than the external codes in space by not launching fixations to relevant, or as a matter of fact, irrelevant regions. Such a looking behaviour can be comparable to cases in which individuals avert their gazes (Glenberg, Schroeder, & Robertson, 1998) or close their eyes (Vredeveldt, Hitch, & Baddeley, 2011) in order to disengage from the environment in the face of cognitive difficulty. Here, we surmise that participants with better visuospatial memory did not feel the "necessity" to rely on the blank locations as their internal memory was sufficient to retrieve the probe word accurately. Thus, they did not look at any regions in a task where moving their eyes could drain cognitive sources further (see Scholz et al., 2018). This interpretation was supported by the fact that participants with better visuospatial memory did better in the memory test in general.

Further, participants with better visuospatial memory looked less at nothing when they saw an incongruent cue. The negative correlation between visuospatial memory and fixations in the relevant quadrant within incongruent cue condition illuminates another dimension of the coordination between internal and external memory. We argue that additional and incongruent visuospatial information made the environment unreliable for a successful memory retrieval. In the event of such spatial interference, participants with better visuospatial memory seemed to ignore any deictic code either attached to words or cues and, turned to internal sources. It appears that unreliability of the external memory was detected as a function of the strength of internal visuospatial memory.

Overall, findings support the integrated memory account (Ferreira et al., 2008; D. C.

Richardson et al., 2009) where internal memory representations and spatial indices which are internalised with eye movements work cooperatively to realise fast and efficient retrieval. On the other hand, results are at odds with the view that looking at nothing is an automatic attempt to access contents of the spatial index (Spivey et al., 2004). If looking at nothing were an automatic behaviour as spatial indexing, all participants would be expected to display the same behaviour regardless of their memory capacity. Rather, results demonstrate that looking at nothing systematically changes not only with the task conditions (e.g., memory demands coming from the task difficulty) (Scholz et al., 2011; Wantz et al., 2015), encoding conditions (e.g., explicit/implicit spatial indexing), or retrieval conditions (e.g., type of retrieval questions, grid arrangement) (Spivey & Geng, 2000) but also cognitive differences between individuals. Coordination between internal and external memory in looking at nothing presents further evidence for the dynamicity account of looking at nothing.

On a larger scale, findings extend the literature showing that the likelihood of cognitive offloading is determined by the abundance of internal sources (Risko & Dunn, 2015). One important aspect here is consciousness. Previous studies showing more frequent cognitive offloading as a consequence of worse internal capacity typically offers offloading as an option to the participants (e.g., Risko & Dunn, 2015). However, looking at nothing is an unintentional and presumably an unconscious behaviour in that participants in our study (and previous looking at nothing studies reviewed above) were never instructed to pay attention to word locations and that they can rely on the environment whenever they encountered retrieval difficulty. Even though, they still used the environment in an intelligent way (Kirsh, 1995) and further, this behaviour was modulated by their internal capacity. Such an unintentional trade-off between internal and external memory might suggest that cognitive offloading to minimise memory load could be a deeply-entrenched but an unconscious memory strategy. That said, consciousness and intentionality were not systematically tested in the current study. Future studies should be designed in a way to investigate whether looking at nothing is a completely unconscious behaviour, or whether we have some kind of control on our "decision" to offload memory work onto the world.

3.6.4 Looking at nothing and memory performance

Results showed that participants who looked at the relevant, blank locations retrieved the probe words more accurately only in congruent cue condition. Accuracy predicted more looks in the relevant quadrant within congruent cue condition as well (see Martarelli et al., 2017; Martarelli & Mast, 2011; Scholz et al., 2014 for looking at nothing occuring in correct trials but not in error trials). Looking at nothing did not predict hit rate in the no cue condition or hit latency in none of the conditions. Thus, we did not present any conclusive evidence that looking at nothing improves memory performance. It can well be argued that Simon-like congruency effect (as in congruent condition) accounts for the enhanced memory rather than fixations in the relevant, blank quadrant. Along with that, eye movements at retrieval were not manipulated in the current study unlike previous studies (e.g., Scholz et al., 2014). Hence, our method to investigate the functionality of gaze position lacks direct causality between looking at nothing and accuracy. Consequently, results reported here cannot distinguish whether participants who looked at nothing were more accurate or participants who were more accurate also looked at nothing. Findings showing partial functionality in the present work should be interpreted cautiously due to methodological limitations.

3.7 Conclusion

Looking at nothing is a unique case in that it demonstrates how the cognitive system can maximize efficiency by spreading the cognitive problem across three domains with the act of looking, the environment with the spatial indices and mental representations in the brain. Our results extended the current literature by shedding further light on the nature of spatial indexing and looking at nothing mechanisms. We provide evidence for automatic and dynamic spatial indexing and a dynamic, efficient looking at nothing behaviour for words. The major contribution of this study is showing a systematic trade-off between internal and external sources driven by individual cognitive differences in order to make the most of environmental opportunities and cognitive capacity. Finally, the current looking at nothing paradigm provides a venue to study the relationship between language and looking at nothing.

Chapter 4

Experiment 2

Simulating Space when Remembering Words:

Role of Memory Load and Word Imageability

4.1 Motivation and Aims

The previous chapter showed that individuals simulate word locations when remembering words. More importantly, results revealed a link between internal visuospatial memory and looking at nothing. This chapter also focuses on the efficient and opportunistic coordination between internal and external memory. Yet, the study described in this chapter addresses the differences between the words to be retrieved rather than the differences between individuals who retrieve the words. In this respect, we move into the domain of language and investigate the effect of different word properties on memory-guided looking behaviour.

4.2 Abstract

People revisit spatial locations of visually encoded information when they are asked to retrieve that information, even when the visual image is no longer present. Such "looking at nothing" during retrieval is likely modulated by memory load (i.e., mental effort to maintain and reconstruct information) and the strength of mental representations. We investigated whether words that are more difficult to remember also lead to more looks to relevant, blank locations. Participants were presented four nouns on a two by two grid. A number of lexico-semantic variables were controlled to form high difficulty and low difficulty noun sets. Results reveal more frequent looks to blank locations during retrieval of high difficulty nouns compared to low difficulty ones. Mixed-effects modelling demonstrates that imagery-related semantic factors (imageability & concreteness) predict looking at nothing during retrieval. Results provide the first direct evidence that looking at nothing is modulated by word difficulty and in particular, word imageability. Overall, the research provides substantial support to the

integrated memory account for linguistic stimuli and looking at nothing as a form of mental imagery.

Highlights

- Remembering less imageable words led to more reliance on the environment during retrieval.
- Experiment 1 was replicated: Participants looked more at relevant, blank locations when remembering single words and worse visuospatial memory led to more reliance on the environment during retrieval.

4.3 Introduction

4.3.1 Memory load and looking at nothing

Under grounded-embodied (Barsalou, 1999; Wilson, 2002) and extended (A. Clark & Chalmers, 1998) views of cognition, human memory exploits available sources in an opportunistic and efficient manner. This is particularly the case in the face of increased cognitive demands (Risko & Gilbert, 2016). Eye movements to “nothing” (i.e., blank locations in space) during memory retrieval is an example of exploitation of external sources to reduce memory load (i.e., the mental effort required for the maintenance and retrieval of information), and to increase memory efficiency (Scholz et al., 2016). In looking at nothing, an integrated memory system attaches spatial information (represented as *spatial indices*) to information that needs to be retrieved during encoding (Ballard et al., 1997; Pylyshyn, 1989; Spivey et al., 2004). When the visual information itself is absent in retrieval, spatial indices trigger eye movements to the blank locations of the previously presented information (Ferreira, Apel, & Henderson, 2008; Richardson & Kirkham, 2004).

The present study aims to specify the conditions under which memory, as an internal faculty, relies on external support via eye movements. Looking at nothing phenomenon presents an appropriate example of how eye movements are employed in such coordination between internal and so-called "external memory". To this end, we examined the mechanisms of looking at nothing by investigating word retrieval with different lexico-semantic properties. To be more precise, we tested whether people rely more on the environmental support by looking at blank locations when remembering words that are more difficult to retrieve from memory due to their lexico-semantic properties.

Previously, eye closure, gaze aversion and other *nonvisual eye movements* that do not involve visual processing but accompany mental operations such as memory retrieval have been shown to be related to cognitive demands (see Salvi & Bowden, 2016 for a review). For instance, people disengage from environmental stimuli by shifting their gaze during challenging memory tasks in order to manage memory load (Doherty-Sneddon & Phelps, 2005; Glenberg, Schroeder, & Robertson, 1998). People also execute eye movements to search for nonvisual information stored in long-term memory and importantly, more frequent eye movements are executed when the task requires a more extensive and conceptually-driven memory search (i.e., verbal memory compared to visuospatial memory) (Ehrlichman & Micic, 2012; Ehrlichman & Barrett, 1983). Ballard et al. (1997) were among the first to suggest that the cognitive system can tap into eye movements at an embodied level in order to minimise memory load. Following a similar line of thought, Spivey and Geng (2000) speculated that people might not look at nothing when the answer in a memory task is salient enough to allow a response before any eye movements are produced. Johansson et al. (2012) also argued that eye movements could serve a supportive role during demanding tasks that involve visuospatial imagery. Similarly, Laeng, Bloem, D'Ascenzo and Tommasi (2014) suggest people attend blank locations only if additional spatial information could make a difference in memory retrieval. Based on these assumptions and previous studies, a correlation can be expected between looks to “now-empty” locations and memory load, where higher load results in more frequent eye movements to blank locations.

Drawing on the potential trade-off between memory load and eye movements, two studies demonstrated that changing cognitive demands coming from task difficulty modulate looking at nothing. First, in Scholz et al. (2011), participants heard four sentences with a visual cue appearing for each sentence in one of the four quadrants across 12 trials. During a retrieval phase, participants’ recognition memory was probed with an auditory statement (correct vs. incorrect) querying a fact from one of the previously presented sentences. Sentences were repeated across the experiment and the proportion of fixations in the relevant, blank quadrant (where the corresponding cue had previously appeared) diminished after the first block as the retrieval task became easier through repetition. Memory load was high in the first block as the information to be retrieved was new and representations were weak. As a consequence, people looked at blank locations. However, the relevant information became familiar by the second and third blocks. Hence, internal memory no longer required an external aid. Consequently, memory load decreased and looking at nothing was not found.

Mental representations were also shown to play a fundamental role in the link between memory load and looking at nothing in a second study showing the effects of decreasing cognitive demands on looking at nothing. Wantz et al. (2015) presented participants with an object in one quadrant of a two by two grid. Memory for the presented objects was probed with a statement about one of the objects as participants looked at a blank screen. The retrieval phase was repeated across five sessions (immediately after the encoding, 5 minutes, 1 hour, 24 hours and 1 week after the encoding). More looks were directed towards the relevant, blank quadrant in the first three sessions compared to looks towards the irrelevant locations. However, looks towards the relevant quadrant were not greater in the latter two sessions, compared to looks towards other locations. In other words, people directed gaze less frequently to spatial locations associated with previously presented objects one day after the original encoding. This suggests that mental representations stabilise over time with repeated retrieval such that revisiting the original locations becomes unnecessary (see also Martarelli & Mast, 2013).

An integrated model of memory (Ferreira et al., 2008; Hoover & Richardson, 2008; D. C. Richardson et al., 2009; Spivey et al., 2004) accounts for the relationship between memory-guided eye movements and memory representations. Under this integrated model, mental representations (internal memory) and environmental sources internalised via eye movements (external memory) work cooperatively. This model is at odds with "world as its own memory" account articulated within the context of radical embodied cognition (e.g., Chemero, 2011). The integrated model acknowledges the existence of mental representations that are "well suited for interfacing with external structures" (Barsalou, 2010). However, "world as its own memory" account rejects the involvement of mental representations in looking at nothing (O'Regan, 1992).

According to the integrated model, representations are integrated and composed of spatial and linguistic input. If one part of the representation (e.g., linguistic) is reactivated through probing, other parts (e.g., spatial information) will be retrieved from memory as well. Task conditions such as repetition can make the linguistic component stronger, which stabilises the mental representations as a whole. In turn, people do not "need" to refer to spatial information for accurate retrieval. Thus, they look less at nothing. In other words, stronger internal memory and stronger mental representations require less environmental support through eye movements. In line with this view, Johansson, Holsanova, and Holmqvist (2011) showed that people with low spatial imagery ability needed more eye movement support when describing a picture from memory using mental imagery. Kumcu and Thompson (2016) also reported less reliance on spatial indices during the retrieval of words among individuals with better visuospatial memory.

There is ample evidence showing that eye movements to blank locations are executed to offload memory work onto the environment during demanding memory tasks (see Risko & Gilbert, 2016 for a review). However, the nature of this behaviour remains elusive: When does the memory system "feel the necessity" to rely on environmental support and when does it turn back to internal memory? Does reliance on the environment change from item to item in a dynamic manner? If so, what type of information drives eye movements to blank locations? We addressed these questions in the current study. More precisely, we investigated whether words that are more difficult to remember also lead to more looks to relevant, blank locations. We hypothesise that fixations to blank locations are more likely to occur during retrieval of more difficult words compared to retrieval of easier words on two accounts: First, retrieval of difficult words impose a higher load on memory, which in turn, may make environmental support more appealing for the opportunistic memory system. Second, if the verbal component of the mental representation is weaker for more difficult words, people may rely on the spatial component more heavily by looking at the original location of the word in order to compensate for the verbal memory deficiency.

If the difficulty of individual items modulates eye movements, then we should expect both increases and decreases in looking percentages from trial to trial within the same session. Evidence for such eye movement behaviour would reveal the ability to switch between internal memory (representations) and so-called "external memory" (spatial indices via eye movements) in a flexible way. There is evidence for the effect of task difficulty on memory-guided eye movements as discussed above. However, we lack direct evidence that looking at nothing is modulated by word difficulty.

4.3.2 Lexico-semantic variables and looking at nothing

Different word properties such as frequency have varying effects on how easily words are remembered. If words that are more difficult to remember lead to more reliance on environmental support via eye movements, the following question arises: Which lexico-semantic variables contribute to looking at nothing? Word properties could affect memory-driven eye movements via two possible channels: Memory load or mental imagery.

Individual properties of a word that make it difficult to remember (e.g., factors such as frequency and age of acquisition) increase memory load (e.g., Collette et al., 2001) and thus, might contribute to the tendency to look at nothing. In this regard, one prediction is that lexico-semantic variables modulate looking at nothing in proportion to their effects on memory performance. Distinctiveness enhances memorability in recognition (Eysenck & Eysenck, 1980; Schulman, 1967). In other words, people are less likely to detect previously seen words if they are not distinctive enough. According to a word difficulty prediction, therefore, variables which make a word less distinct and thus, more difficult to remember will also contribute to looks to blank locations during retrieval.

Many word properties play a role in verbal recognition memory through distinctiveness. For example, worse recognition performance has been evidenced for more frequent (Glanzer & Bowles, 1976), more available words (i.e., words that come to mind easily) (Rubin, 1983), early-acquired words (Dewhurst, Hitch, & Barry, 1998; but see Coltheart & Winograd, 1986) and words that have more orthographically similar neighbours (e.g., “book” – “hook”, “cook”, “crook” etc.) (Cortese, Khanna, & Hacker, 2010; Cortese, McCarty, & Schock, 2015). Longer words are typically regarded as more distinct. However, more hits (i.e., correctly identifying previously seen words as old words) and fewer false alarms (i.e., identifying new words as previously seen) were reported for shorter words, suggesting that longer words tax the memory system by imposing more load (Cortese et al., 2010, 2015). In a typical recognition memory paradigm, words are presented visually and hence, encoded and retrieved in written form. As a result, phonological effects have not been demonstrated. For example, phonological similarity (i.e., having more neighbours that sound similar) does not appear to predict recognition memory accuracy as opposed to orthographic similarity (e.g., Cortese et al., 2010, 2015). Pronounceability was reported to have limited effect on recall (Rubin & Friendly, 1986) and its effect on recognition memory is not clear.

Imageability has a critical role in verbal memory (see Paivio, 1991; Schwanenflugel, 1991 for reviews). Imageability is defined as the extent to which a word evokes a mental image (Paivio, Yuille, & Madigan, 1968). For example, "apple" is a highly-imageable word in that its meaning can quickly bring a salient picture to mind, which would be similar for most people. Whereas, the same cannot be said for low-imageable words such as "offer" or "coincidence". Although these words can also stimulate images to a certain degree, they would not be as strong as those associated with high-imageable words. It is well established that words associated with perceptually salient, highly imageable objects/concepts are better remembered than those associated with low imageable objects/concepts (see Marschark & Cornoldi, 1991 for a review). Imageability was shown to be one of the strongest predictors of recognition memory (e.g., Cortese et al., 2010, 2015) and recall (Rubin & Friendly, 1986) relative to other variables.

One prediction tested here is that imageability modulates looking at nothing due to its contribution to the mental image of the target word's referent. Decades of evidence has demonstrated that eye movements are instrumental in mental imagery processes (see Mast & Kosslyn, 2002 for a review). For instance, Noton and Stark (1971) showed that eye movements during imagery are similar to the movements during perception (scanpath theory). Specifically, people simulate perception during imagery by re-enacting the eye movements that are executed during viewing (Altmann, 2004; Johansson et al., 2012; Laeng & Teodorescu, 2002; but see Johansson, Holsanova, & Holmqvist, 2006). Further, eye movements in mental imagery appear to support the image generation process (Johansson et al., 2012; Laeng et al., 2014). For instance, the degree of similarity in scanpaths between perception and imagery predicts the accuracy of memory for the visual scene (Laeng & Teodorescu, 2002).

Previously encoded visual or verbal information is "recreated" without any visual stimulus when people attend to blank locations during retrieval. Thus, looking at nothing involves visuospatial mental imagery by nature. Low-imageable words are expected to have weaker mental images as opposed to high-imageable words. In the face of weak mental images, people could rely more on external support by looking at blank locations to meet the imagery deficit. Alternatively, participants might treat the words as "picture-like" orthographic units when remembering them on a blank screen. In such a case, mental images are expected to reflect the physical, perceptual properties of the words (e.g., number and shape of letters) rather than conceptual elements activated by word meanings (see Hunt & Elliot, 1980). Under this prediction, orthographic properties; namely, word length, number of syllables and orthographical similarity would be expected to regulate fixations to the blank location.

In light of the research discussed above, we selected ten variables to control the words in the current study (imageability, concreteness[3], context availability, pronounceability, age of acquisition, frequency, syllable length, length in letters, phonological and orthographic similarity). Mixed-effects models were fit to reveal the predictors of looking at nothing. It is important to note that imageability is a crucial predictor in both word difficulty and mental imagery predictions. Thus, models were fit for the variables predicting memory performance as well. If memory-guided eye movements are modulated mainly by word difficulty and

[3] Imageability and concreteness are highly correlated, exhibit similar advantages in memory (see also Richardson, 1975) and thus, used interchangeably in many studies (e.g., Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Nittono, Suehiro, & Hori, 2002; Reilly & Kean, 2007). Along with that, we controlled the words on concreteness in addition to imageability due to conceptual differences between them (see Vigliocco, Vinson, Druks, Barber, & Cappa, 2011 for a discussion). Concreteness is typically defined as the extent to which a word can be experienced by senses.

memory load, predictors of memory performance should also predict looking at nothing. If mental imagery, in particular, modulates looks to blank locations, then imageability should stand out as a critical predictor of looking at nothing. If mental images corresponding to words are based on orthographic properties, word length (in letters and syllables) and orthographical similarity rather than imageability should play a role in eye movements to blank locations.

4.3.3 Spatial interference between encoding and retrieval

Finally, we aimed to follow the experimental design in Experiment 1 (Kumcu & Thompson, 2016) for consistency and comparison. Thus, participants were presented black dots as unrelated visual cues between encoding and retrieval phases. Cues were either congruent (shown in the same location as to the original location of the probe word) or incongruent (shown in a diagonal location as to the original location of the probe word) in addition to a "pure" looking at nothing condition without any cue. The cueing condition was designed to guide participants' attention and eye movements to the location of the information held in memory (congruent cue) or away from it (incongruent cue) before retrieval.

There is evidence that additional visual processing within the looking at nothing paradigm has consequences both on memory performance and eye movements. For example, in Scholz, Klichowicz and Krems (2017), participants were asked to judge the truth of a sentence they had encoded in a grid location. At the same time, they were asked to attend a visual tracking task (L. E. Thomas & Lleras, 2009). In this task, random string of digits from 0 to 9 appeared on the screen and participants had to press a button whenever the digits appeared. Importantly, digits always appeared in the same location of the grid in a trial; that is, either congruent or incongruent locations with the location associated with the sentence to be retrieved. In one condition (overt attention), participants were asked to gaze freely as the visual tracking task occurred on the retrieval screen. In the other (covert attention), participants were asked to fixate on the centre and respond when the digits appeared. Participants were less accurate when the digit appeared in the incongruent locations compared to the congruent locations both under overt or covert attention conditions. Similarly, Kumcu and Thompson (2016) showed that a visual cue shown between the encoding and retrieval stages which is congruent with the location of the to-be-remembered word reinforces the spatial index of the word and thus, amplifies looking at nothing. On the other hand, an incongruent visual cue interferes with spatial indexing of the probe word and leads to the disruption of looks to relevant, blank locations.

In the present study, we investigated how spatial cues modulate the link between retrieval difficulty due to lexico-semantic variables and looking behaviour/memory performance. In particular, we tested how imageability affects looking at nothing in congruent and no cue conditions respectively. If mental imagery is a reinstatement of previous perceptions in the absence of any stimulus (Hebb, 1968; Kosslyn, Thompson, & Ganis, 2006), the effect of imageability on looking at nothing should be stronger in a no cue condition in comparison to a congruent cue condition. There is overwhelming empirical evidence that actual visual perception and visual mental imagery share common mechanisms and influence each other (Cichy et al., 2012; Ganis et al., 2004; Kosslyn et al., 1997; Perky, 1910). Hence, visual perception of a cue could interfere with the generation of a mental image invoked with words under congruent cue condition. In turn, this could attenuate the effect of word imageability on looking at nothing. Whereas, word imageability could modulate looking behaviour under no cue condition; that is, when there is no visual information to interfere between encoding and retrieval phases.

4.4 Method

4.4.1 Participants

The experiment was carried out with forty-eight students at the University of Birmingham (nine males; M_{age} = 20.06, *SD* = 2.30, range: 18 - 29). 75% of them were psychology students. All participants were monolingual native speakers of British English as determined with the Language History Questionnaire version 2.0 (Li, Zhang, Tsai, & Puls, 2013). Participants reported normal or corrected-to-normal vision, no speech or hearing difficulties and no history of any neurological disorder. They received either £10 (*n* = 27) or course credit (*n* = 21) for participation. All participants were fully informed about the details of the experimental procedure and gave written consent. Post-experiment debriefing revealed that all participants were naïve to the purpose of the experiment. No participant was replaced.

4.4.2 Materials

There were 180 trials involving 810 unique nouns in total. All words were drawn from the extensions of Paivio, Yuille and Madigan norms for 925 nouns (J. M. Clark & Paivio, 2004). The word pool was filtered to exclude words shorter than three letters and longer than 11 letters.

Trials were evenly divided into two (*n* = 90) as high difficulty and low difficulty word groups based on the mean imageability of the whole set (4.99). It is not viable to manipulate one dimension by holding others constant due to intercorrelations between the variables. Thus, high difficulty words were less imageable, more abstract, less available, less pronounceable, learnt later in life, longer (both in number of letters and syllables) and had less phonologic and orthographic similarity with other words in the language (see Table 4.1).

Both high and low difficulty groups were further divided into yes and no trials (n = 45). Probe words in the yes trials were among the four study words in the encoding phase, whereas a different, not seen, word was probed in the “no trials”. There were no significant differences in any of the variables between yes and no groups within high and low difficulty word groups (all ps > .05).

Words were then grouped into smaller trial sets of four (yes trials) and five words (no trials). Words within sets were matched on all variables (all SDs < 2.00) both in the yes and no trials. Words were further controlled such that no word started with the same letter or had any semantic relationship with any other word in the set. Monosyllabic, disyllabic and trisyllabic words were evenly distributed [e.g., (3, 3, 3, 3), (1, 2, 1, 2) or (3, 2, 3, 2, 3) etc.].

The word in each trial set with median imageability was selected as the probe leaving the others as distractors. Welch’s t-tests revealed no significant differences between the probe and distractor words in any of the variables or in any of the four sub groups (i.e., high difficulty yes, low difficulty yes, high difficulty no, low difficulty no) (all ps > .05). Thus, any word among the four or five words in each trial set was as likely to be remembered as any other word in the same set.

Finally, we formed 180 unique mathematical equations [e.g., (2*3) - (2+3) = 1] to present as memory interference between encoding and retrieval phases (see Conway & Engle, 1996 for a similar design). Half of the equations were correct. Incorrect equations were further divided into two equal groups: The results were either plus or minus one of the correct result.

Table 4.1 Differences in lexico-semantic variables between high and low difficulty words shown as mean values, standard deviations in parentheses and Welch's t-test statistics

Variable	High difficulty words	Low difficulty words	*t*	*p*
Imageability	3.78 (0.88)	6.19 (0.43)	49.55	< .0001
Concreteness	3.55 (1.45)	6.51 (0.65)	37.41	< .0001
Length in letters	7.45 (1.88)	6.13 (1.78)	-10.30	< .0001
Number of syllables	2.59 (0.90)	1.85 (0.77)	-12.64	< .0001
Orthographic similarity	2.89 (0.52)	3.15 (0.73)	5.63	< .0001
Phonological similarity	2.80 (0.73)	3.32 (1.04)	8.22	< .0001
Pronounceability	6.23 (0.63)	6.53 (0.43)	8.12	< .0001
Age of acquisition	4.96 (0.89)	3.64 (1.05)	-19.29	< .0001
Availability	2.07 (0.80)	2.28 (0.78)	3.68	< .001
Frequency	1.10 (0.70)	1.16 (0.66)	1.09	.28

4.4.3 Apparatus

Stimuli were presented on a TFT LCD 22-inch widescreen monitor operating at 60 Hz with a resolution of 1680 x 1050 pixels (501.7 mm x 337.4 mm). The monitor was placed 640 mm in front of the participant. A chin and forehead rest was used to reduce head movements. Participants' eye movements were monitored using SR EyeLink® 1000 (sampling rate: 1000 Hz, spatial resolution < 0.5°, http://sr-research.com/eyelink1000.html). Viewing was binocular but only the left eye was monitored. Auditory material was produced by a native female speaker of British English in a sound attenuated room and recorded using Audacity (version 2.1.10, https://www.audacityteam.org). Participants responded (yes/no they had seen the word) by pressing one of two keys on a standard keyboard. Eye movement data were extracted using the SR EyeLink Data Viewer (version 2.4.0.198, https://www.sr-research.com/data-viewer/). No drift or blink correction procedure was applied.

Data were analysed and visualised in R programming language and environment (R Core Team, 2017). Mixed-effects models were constructed with *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). Significance values of the coefficients in models were computed based on the t-distribution using the Satterthwaite approximation with *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015).

4.4.4 Procedure

We followed the procedure in Experiment 1 (Kumcu & Thompson, 2016). Eye tracking started with a standard nine-point calibration and validation, which confirmed high data quality (average calibration error < 1° and maximum calibration error < 1.50°). As spelled out in detail below, each trial was composed of five consecutive phases: (1) fixation, (2) encoding, (3) cueing, (4) interference and (5) retrieval (See Figure 4.1). The task was to decide whether an auditorily presented word had appeared before or not (i.e., yes/no verbal recognition memory test). As soon as the participants made yes/no judgement by hitting one of the response buttons, the trial ended, and a new encoding phase began.

(1) Fixation: A fixation cross appeared at the centre of the screen for 500 ms. **(2) Encoding**: Participants were presented four words in capital letters on a 2 x 2 grid for 1800 ms. Words (Times New Roman, font size = 40) were centrally placed in rectangular boxes (300 x 85 in pixels, 8° x 2.4° of visual angle). Word difficulty was a within-subjects variable and all participants saw the high and low difficulty words. **(3) Cueing**: A flashing black dot appeared in cue trials for 1000 ms either in the same (congruent cue) or in the diagonal quadrant (incongruent cue) as the original location of the probe word in the encoding phase. There was also a third condition where no cue was presented between encoding and interference. Cue condition was a within-subjects variable and three cue conditions were randomly presented in a session. That said, an equal number of random participants ($n = 16$) saw the same probe word with a congruent cue, an incongruent cue or without any cue. **(4) Interference**: Participants were presented a mathematical equation and asked to identify whether the equation was correct or not within 20,000 ms (or they timed-out). **(5) Retrieval**: The probe word was auditorily presented as participants looked at the blank grid with empty boxes. There was a 500 ms gap between the presentation of the blank retrieval screen and the presentation of the probe word. Participants were asked to make an unspeeded yes/no judgement to determine whether they had seen the probe word among the four words shown in the encoding phase within 20,000 ms (or they timed-out).

The order of trials and equations were fully randomised independent of each other. The location of all words in all conditions was counterbalanced with Latin Square design to control gaze biases so that each word appeared an equal number of times in each location of the grid. The experiment was divided into four equal blocks with 45 trials in each block and there was a short pause between blocks. A typical session lasted approximately 100 minutes, including consent and setting up the eye tracker. Overall accuracy in interference equations and in the recognition memory test for words were 88% and 78% respectively, suggesting that participants attended to the task with high concentration.

Following the experiment, a computerized version of the Corsi block-tapping task (Corsi, 1972) operated on PEBL (Psychology Experiment Building Language, version 0.13, test battery version 0.7, http://pebl.org) (Mueller & Piper, 2014) was used to measure visuospatial short-term memory and Gordon Test of Visual Image Control (Gordon, 1949) was administered to measure subjective ability with regard to manipulation of mental images.

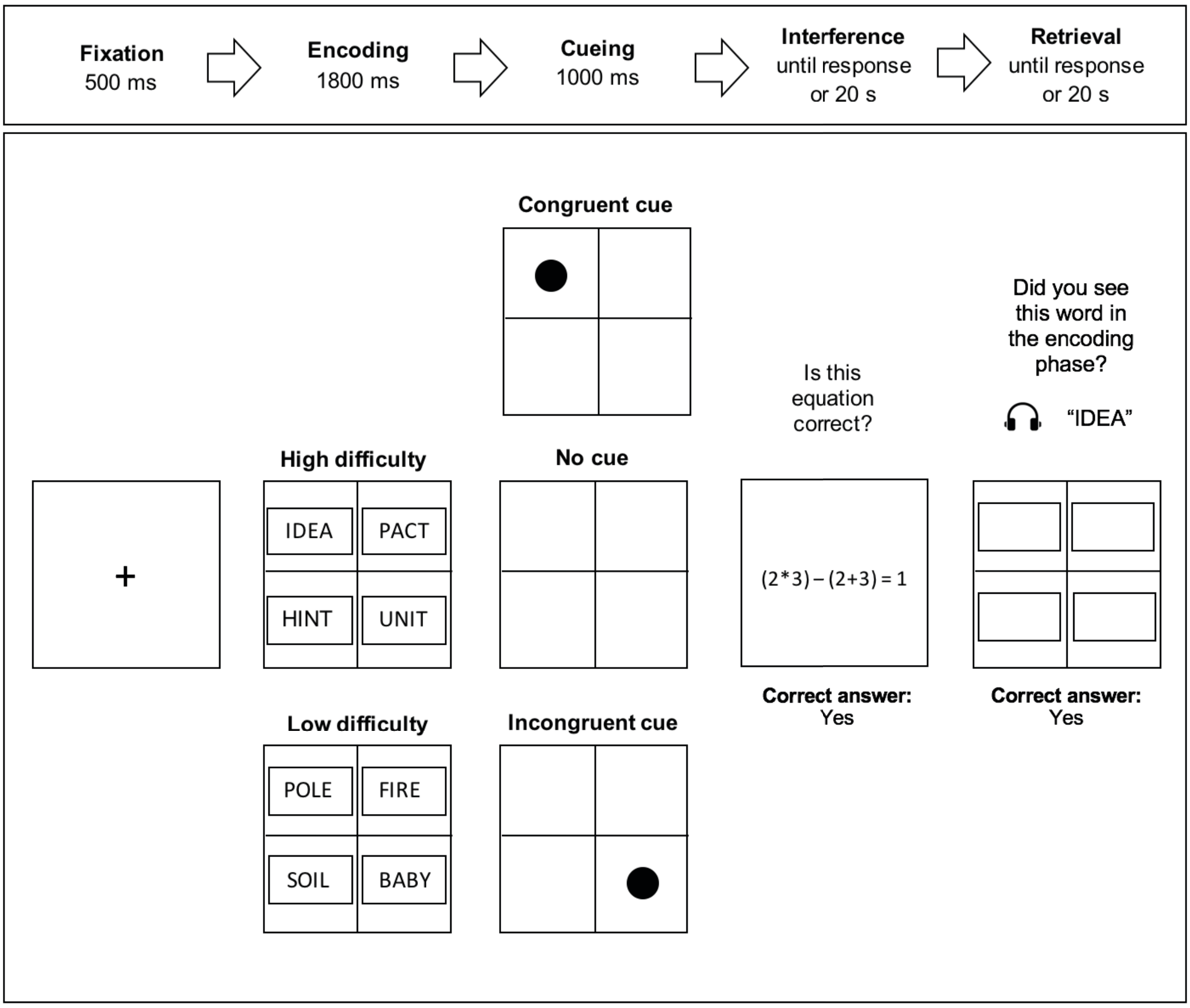

Figure 4.1 A schematic illustration of the temporal order of events in an example trial showing high and low difficulty word conditions and three different cue conditions. In this example, the relevant quadrant is the top left location, where the probe word (i.e., IDEA) appears.

4.5 Results

4.5.1 Measures

Results were analysed in two parts as memory performance and looking behaviour.

Memory performance: Hit rate, hit latency and correct rejection rate were used as measures of memory performance. Hit rate was the proportion of yes trials to which the participants correctly responded yes. Correct rejection rate was the proportion of no trials to which the participants correctly responded no. Hit latency was the time in milliseconds between the onset of auditory presentation of the probe word and correct keyboard response. Participants were not instructed to make speeded response in the current paradigm. Nevertheless, hit latencies were reported to verify and complement the indicators of memory performance based on accuracy.

Looking behaviour: Fixation percentage was used as the main gaze measure and dependent variable as in previous looking at nothing studies discussed above (e.g., Wantz et al., 2015). Fixation percentage (or fixation frequency) is the percentage of fixations in a trial falling within a particular interest area in proportion to total fixations in a trial. Thus, it was computed by dividing the number of fixations on each quadrant to the total number of fixations during the retrieval phase (see Wenzel, Golenia, & Blankertz, 2016 for a similar computation and use of fixation frequency).

Proportion of fixations was selected as the main indicator of looking behaviour on two grounds: First, it is immune to differences in durations. To be more precise, fixation percentage was considered appropriate particularly for comparing two different conditions (high and low difficulty words) with varying trial durations due to differences in word length between high and low difficulty words. Second, we assumed that fixations rather than the time spent on particular region (i.e., dwell time per quadrant) is important for the link between memory and eye movements. Fixation-based measures are reliable indicators of memory load and attention in a given location (e.g., Just & Carpenter, 1980; Meghanathan, van Leeuwen, & Nikolaev, 2015). Hence, we preferred fixation percentage over dwell time percentage as a more refined indicator of looking at nothing[4]. Accordingly, we expected that participants would fixate on the relevant quadrant to derive support from the environment.

Four rectangular interest areas corresponding to the quadrants were identified. All interest areas were of the same size (502 x 368 in pixels, 13.4° x 10.6° of visual angle). They framed the rectangular boxes that words were presented in (see Figure 4.1) and were not contiguous. Proportion of fixations accrued on the interest areas during the retrieval phase (from the onset of auditory presentation of the probe word until the participant's response) were calculated. Fixations were a minimum duration of 40 ms. First fixations and fixations outside the interest areas (8.62%) were omitted. Only hits (i.e., correct responses) in yes trials were included in the fixation analyses. Fixation percentages allocated to the three quadrants that did not contain the target probe word were averaged into one and analysed against the relevant quadrant in which the probe word was seen.

4.5.2 Mixed-effects modelling

Data were analysed using linear and binomial logit mixed-effects modelling. Visual inspections of residual plots did not reveal any obvious deviations from homoscedasticity or linearity. Linear models were fit for continuous target variables (hit latency and fixation percentage). Binomial models were fit for categorical target variables (hit rate and correct rejection rate) and with *bobyqa* optimiser to prevent non-convergence. Participants and items were treated as

[4] The same analyses were performed with dwell time percentages as well and findings were consistent with the analyses based on fixation percentages reported here.

random effects to explain by-participant and by-items variation (Baayen, Davidson, & Bates, 2008).

We started fitting models by building the random effects structure and followed a maximal approach. That is, random effects were included as both random intercepts and correlated random slopes (random variations) as long as they converged and were justified by the data (Barr, Levy, Scheepers, & Tily, 2013). Random intercepts and slopes were included even if they did not improve the model fit in order to control for possible dependence due to repeated measures or order effects. Random effects structure was simplified step by step as per the magnitude of the contribution of a random effect to the explanation of the variation in the data. That is, the random effect with the weakest contribution was dropped first and if necessary, the structure was further reduced accordingly.

Two approaches were adopted when building the fixed effects structure. Contribution of a fixed effect was investigated by comparing a full model containing the effect in question against a reduced model in which only that effect was removed, or a null model without any fixed effects. Compared models had the same random effects structure (Winter, 2013). Best-fit model was specified by starting with a full model which included all fixed effects and their interactions (Bates et al., 2015). The full model was then reduced systematically in each step until the null model. Models were then compared using *anova* function to identify the model offering the best-fit by Akaike information criterion (AIC) and Bayesian information criterion (BIC), where lower is better in both cases (Hilbe, 2011) (see Appendix for the outputs of the best-fit models).

4.5.3 Factor analysis

The effect of lexico-semantic predictors on looking at nothing and memory performance were examined. Thus, we conceptualised fixation percentage in blank location, hit rate, hit latency and correct rejection rate as a function of length in letters, syllable length, orthographic similarity, phonological similarity, age of acquisition, frequency, availability, pronounceability, imageability and concreteness.

However, diagnostic tests indicated collinearity between the predictors as identified with the correlation matrix, variance inflation factors (VIF) (M = 3.60, range = 2.42 - 5.13) and kappa (96.34). In order to address collinearity, we performed exploratory factor analysis and clustered the variables in components. Results of Bartlett's test of sphericity, $\chi^2(45)$ = 17011.77, p < .0001 and Kaiser-Meyer-Olkin measure of sampling adequacy (.78) supported the existence of factors within the data. Hence, we proceeded to conduct the factor analysis using a principal component analysis extraction method with an orthogonal (varimax) rotation method. Both Kaiser's criterion and the scree test criterion indicated the presence of three factors in our data. This conclusion was also supported by the percentage of variance criterion (Hair, Black, Babin, & Anderson, 2009), which suggests that all retained factors should account for at least 60% of the total variance. The three-factor solution in our analysis explained 78.24% of the variance in the data. Communalities of all predictors were above .70 (Mean communality = 78.23; see Table 4.2) suggesting that all measures were adequately accounted for by the three-factor solution. As suggested by (Hair et al., 2009), only factor loadings above .40 (or below -.40) were considered to meet the minimal level for interpretation of factor structure. Factor loadings did not have substantial loadings on other factors and they showed particularly clean clustering (except for age of acquisition). Further, the factors themselves had substantial loadings only for those variables, thus other variables did not load on these factors.

Three factors were interpreted as follows based on the loadings (see Table 4.2): **(1) Imagery:** Imageability and concreteness. **(2) Lexical:** Age of acquisition, frequency, availability and pronounceability. **(3) Length & similarity**: Length in letters, syllable length, orthographic similarity and phonological similarity. Age of acquisition seemed to contribute to all three factors to a certain degree. It was therefore considered within the lexical factor on theoretical grounds, its loading and in line with previous factor analyses (e.g., J. M. Clark & Paivio, 2004) VIF of the factors were below one and thus, below our threshold of two.

Regression scores calculated for the three factors were employed both as predictors and random slopes in the subsequent linear mixed-effects multiple regression models. Regression scores were additionally recalculated for each subset in each analysis. As expected, factor analyses extracted similar factor loadings and produced the same three-factor solutions.

Table 4.2 Varimax rotated factor-loadings and communalities of the predictors

	Factor			
Predictors	Imagery	Lexical	Length & similarity	h^2
Imageability	**0.91**	0.17	0.25	0.93
Concreteness	**0.92**	-0.04	0.18	0.89
Frequency	-0.05	**0.91**	0.00	0.82
Availability	0.01	**0.81**	0.29	0.74
Pronounceability	0.20	**0.79**	0.33	0.76
Age of acquisition	-0.48	**-0.61**	-0.37	0.74
Phonological similarity	0.14	0.20	**0.89**	0.85
Syllable length	-0.18	-0.17	**-0.87**	0.82
Length in letters	-0.20	-0.13	**-0.84**	0.77
Orthographic similarity	0.22	0.30	**0.73**	0.68
Factor statistics				
Eigenvalue	2.10	2.67	3.22	7.99
Variance (%)	21.03	26.73	32.16	79.93

Bolded numbers indicate the groupings. Eigenvalues and percentage of variance are after rotation. h^2 = communality

4.5.4 Memory performance

Hit rate

First, we analysed whether there was a difference in hit rate across congruent and incongruent cue conditions. The fixed effect was cue location with two levels (congruent and incongruent cue). Imagery and length & similarity factors were added as random slopes into participants. Imagery, lexical and length & similarity factors were added as random slopes into items. Cue location did not improve the model fit when compared against a null model; $\chi^2(1) = 1.16$, $p = .28$. In other words, participants retrieved the probe words in incongruent cue condition (mean hit rate = 79%) as accurately as congruent cue condition (mean hit rate = 77%). Cue location did not improve the model fit either when no cue condition (mean hit rate = 78%) was included; $\chi^2(2) = 1.18$, $p = .55$.

Second, we examined lexico-semantic variables modulating hit rate. As reported above, we did not find any differences in hit rate across cue conditions. Thus, mixed-effects models including all cue conditions were fit. All factors (i.e., imagery, lexical and length & similarity factors) were added as random slopes both into participants and items. Imagery factor improved the model fit significantly; $\chi^2(1) = 13.20$, $p = .0003$. Length & similarity factor contributed to the model with even higher magnitude; $\chi^2(1) = 24.49$, $p < .0001$. Whereas, lexical factor was not predictive of hit rate; $\chi^2(1) = 0.02$, $p = .90$. The best-fit model converged with length & similarity and imagery factors; $\chi^2(1) = 13.19$, $p = .0003$. This model was supported by a large AIC difference of 11.2 and a BIC difference of 4.8 compared to the next best model converged with length & similarity factor only. Participants were more accurate when retrieving high imageable & concrete; $\beta = 0.18$, $z = 3.96$, $p < .0001$ and shorter & less similar words; $\beta = 0.24$, $z = 4.86$, $p < .0001$.

Length & similarity included four different variables (phonological similarity, orthographic similarity, length in letters and syllable length) which might have contradicting effects on the memory performance. We therefore fitted simpler models in order to identify the individual effects of the variables on hit rate within length & similarity factor. In order to avoid collinearity, we selected length in letters from the word length set and orthographic similarity from the similarity set as fixed effects (VIF < 2). Models including word length and orthographic similarity as random slopes within items and word length within participants indicated that hit rate was predicted by word length; $\chi^2(1) = 23.89$, $p < .0001$ but not orthographic similarity; $\chi^2(1) = 0.12$, $p = .73$. Participants were more accurate when retrieving shorter words; $\beta = -0.12$, $z = 5.79$, $p < .0001$. Models with syllable length and phonological similarity did not change the results.

Hit latency

Linear mixed-effects models were fit to identify any difference in hit latency between cue conditions. Imagery and length & similarity factors were added as random slopes into participants. Imagery, lexical and length & similarity factors were added as random slopes into items. As in hit rate, likelihood tests indicated that there was no difference in hit latency between congruent (mean hit latency = 1978.22 ms) or incongruent (mean hit latency = 2057.67 ms) cue conditions; $\chi^2(1) = 2.16$, $p = .14$. Results did not change when no cue condition (mean hit latency = 2027.86 ms) was included; $\chi^2(2) = 2.04$, $p = .36$.

Next, we investigated the effect of lexico-semantic factors on hit latency. Imagery, lexical and length & similarity factors were added as random slopes into participants. Imagery and length & similarity factors were added as random slopes into items. All three factors, that is, lexical factor; $\chi^2(1) = 4.62$, $p = .03$, imagery factor; $\chi^2(1) = 6.14$, $p = .01$ and with a considerably higher magnitude, length & similarity factor; $\chi^2(1) = 34.29$, $p < .0001$ predicted hit latency.

Thus, the best-fit model converged with all three factors as the fixed effects; $\chi^2(1) = 4.34$, $p = .04$. Participants were faster to retrieve high imageable & concrete; $\beta = -48.66$, $t = 2.38$, $p = .02$ and shorter & less similar words; $\beta = -83.24$, $t = 4.73$, $p < .0001$. They were slower to retrieve more frequent, more available, more pronounceable words which were learned earlier in life; $\beta = 41.49$, $t = 2.13$, $p = .04$.

As in hit rate, simpler models with word length and orthographic similarity demonstrated that hit latency was predicted by word length; $\chi^2(1) = 8.57$ $p = .003$ but not orthographic similarity; $\chi^2(1) = 0.49$, $p = .49$. Participants were faster to retrieve shorter words; $\beta = 44.5$, $t = 4.78$, $p < .0001$. Models with syllable length and phonological similarity did not change the results. Correlation between variables within lexical factor (i.e., frequency, availability, pronounceability and age of acquisition) did not allow us to investigate their individual effects on hit latency due to high collinearity (VIF > 2).

Correct rejection rate

Correct rejection rate was the proportion of “no trials” to which the participants correctly responded no. Visual cues in yes trials were located according to the location of the probe words at encoding; whereas, a different, not seen, word was probed in no trials. Thus, “no trials” were not presented with different cue conditions and there is necessarily no effect of cue for these trials. As a result, only the effect of lexico-semantic variables on correct rejections was investigated.

Imagery, lexical and length & similarity factors were added as random slopes into items. Participants were added as a random intercept as the random-slope model did not converge. None of the lexico-semantic variables predicted correct rejection rate: [length & similarity factor; $\chi^2(1) = 3.23$, $p = .07$, imagery factor; $\chi^2(1) = 0.13$, $p = .72$, lexical factor; $\chi^2(1) = 0.23$,

p = .63]. Word length and orthographic similarity as raw values rather than regression scores of the length & similarity factor did not predict correct rejection rate either.

4.5.5 Looking behaviour

Looking at nothing

We first analysed whether participants looked at nothing during memory retrieval. In other words, we investigated whether there were more looks to relevant, blank locations where probe words were shown at the encoding stage compared to irrelevant, blank locations.

The target variable was fixation percentage in correctly answered yes trials. The fixed effect was quadrant with two levels (relevant and irrelevant quadrant). Imagery, lexical and length & similarity factors were added as random slopes into participants and items.

Quadrant improved the model fit when all cue conditions were factored in; $\chi^2(1) = 8.60$, $p = .003$. Models with a single cue condition indicated that quadrant improved the model fit in congruent; $\chi^2(1) = 5.70$, $p = .02$ and in no cue conditions; $\chi^2(1) = 5.39$, $p = .02$ but not in incongruent cue condition; $\chi^2(1) = 0.13$, $p = .72$. That is, participants looked more at the relevant location in congruent; $\beta = 0.02$, $t = 2.39$, $p = .02$ and no cue conditions; $\beta = 0.02$, $t = 2.32$, $p = .02$. However, they did not look at nothing when they were shown a visual cue that was incongruent with the original location of the probe word during encoding. Hence, we analysed the effect of word difficulty and lexico-semantic predictors on looking at nothing in congruent and no cue conditions.

Effect of word difficulty on looking at nothing

A further sub-analysis was performed on the two conditions with overall evidence of looking at nothing behaviour (congruent and no cue conditions), in order to determine the role of word difficulty. Linear mixed-effects models with high and low difficulty word sets were fit separately. The target variable was fixation percentage in correctly answered yes trials. Fixed effect was quadrant with two levels (relevant and irrelevant quadrant). All lexico-semantic factors were included into participants and items as random slopes.

Congruent cue condition: Quadrant improved the model fit for high difficulty; $\chi^2(1) = 7.00$, $p = .008$ but not low difficulty word set; $\chi^2(1) = 0.59$, $p = .44$. Participants looked more at the relevant quadrant than the irrelevant quadrant only when retrieving more difficult words; $\beta = 0.04$, $t = 2.65$, $p = .008$ (see Figure 4.2).

No cue condition: Quadrant improved the model fit in high difficulty words set with a higher magnitude than the congruent cue condition; $\chi^2(1) = 9.59$, $p = .002$. Quadrant did not improve the model fit in low difficulty words set; $\chi^2(1) = 0.08$, $p = .77$. Participants looked more at the relevant quadrant than the irrelevant quadrant when retrieving more difficult words; $\beta = 0.04$, $t = 3.10$, $p = .002$ (see Figure 4.2).

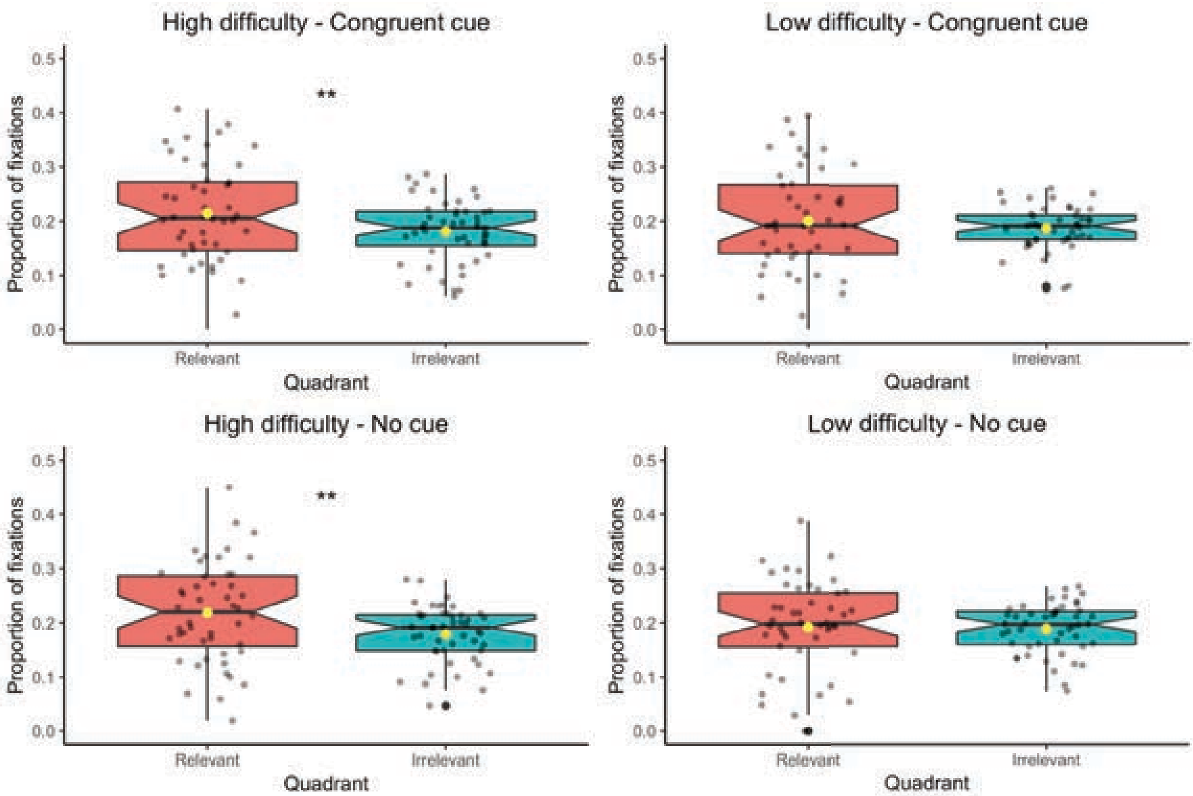

Figure 4.2 Proportion of fixations across relevant and irrelevant quadrants as participants retrieved high difficulty and low difficulty words in congruent and no cue conditions. Values on the y axis correspond to percentages (e.g., 0.1 = 10%). Notched box plots show median (horizontal line), mean (yellow dot), 95% confidence interval of the median (notch), interquartile range (the box), the first and the third quartiles (lower and upper ends of the box) and ranges (vertical line). Grey dots represent data points. ** $p \leq .01$

Hit rate, hit latency and correct rejection rate between high and low difficulty words were compared with the aim of confirming the reliability of the difficulty manipulation. As expected, participants were more accurate; $\beta = 0.36$, $z = 4.51$, $p < .0001$ and faster; $\beta = -118.58$ $t = 3.37$, $p = .0008$ when retrieving the low difficulty words. Further, low difficulty words were subjected to less false alarms; $\beta = 0.32$, $z = 2.62$, $p = .009$.

Lexico-semantic predictors of looking at nothing

Next, we aimed to disambiguate the variables composing word difficulty. Effect of lexico-semantic predictors on looking at nothing behaviour was investigated. The target variable was fixation percentage in the relevant quadrant in correctly answered yes trials. Imagery and length & similarity factor were added as random slopes into participants. Imagery factor was added as random slopes into items.

Congruent cue condition: Imagery factor; $\chi^2(1) = 3.21$, $p = .07$, length & similarity factor; $\chi^2(1) = 1.27$, $p = .26$ or lexical factor; $\chi^2(1) = 0.24$, $p = .62$ did not significantly predict fixation percentage in the relevant quadrant. Along with that, the best-fit model explaining the data was fit with imagery factor; $\beta = -0.02$, $t = 1.76$, $p = .08$ based on AIC (0.74) and BIC differences (5.75) as to the next best model converged with imagery and length & similarity factors.

No cue condition: Imagery factor predicted fixation percentage in the relevant quadrant; $\chi^2(1) = 4.31$, $p = .04$. Length & similarity; $\chi^2(1) = 1.37$, $p = .24$ or lexical factor; $\chi^2(1) = 0.69$, $p = .41$ were not significant predictors of looking at nothing. As a result, the best-fit model explaining the data was the one with imagery as the fixed effect; $\chi^2(1) = 4.06$, $p = .04$. Higher imageability and concreteness predicted less fixations in the relevant quadrant; $\beta = -0.02$, $t = 2.03$, $p = .04$.

4.5.6 Additional analyses

Functionality of looking at nothing

The current experiment was not designed to test the functionality of looking behaviour in memory. Nevertheless, we examined whether memory performance was predicted by the proportion of fixations in the relevant, blank locations. Imagery factor was added as random slopes into participants. Imagery and length & similarity factors were added as random slopes into items. Looks to relevant, blank locations did not predict hit rate (congruent cue; $\beta = 0.30$, $z = 1.24$, $p = .22$, no cue; $\beta = 0.02$, $z = 0.09$, $p = .93$) or hit latency (congruent cue; $\beta = -64.77$,

$t = 0.70$, $p = .49$, no cue; $\beta = 85.31$, $t = 0.78$, $p = .44$). Models fit with high difficulty or low difficulty words only did not change the results.

Mental imagery control and looking at nothing

Additionally, we investigated the correlation between subjective measures of mental imagery control and looking at nothing. There was a positive correlation between mental imagery control and fixation in the relevant quadrant in no cue condition; r_s (46) = .34, $p = .02$ but not in congruent; r_s (46) = .23, $p = .11$ or incongruent cue condition; r_s (46) = .04, $p = .81$. Participants who reported to control their mental imagery better used space more frequently in the memory task when there was no visual cue between encoding and retrieval stages.

Visuospatial memory and looking at nothing

The current experiment did not specifically address the link between visuospatial memory and looking at nothing. Along with that, we investigated whether the effect of visuospatial memory on looking behaviour is robust enough to replicate. Results in Experiment 1 were replicated. There was a significant, negative correlation between visuospatial memory capacity and fixation percentage in the irrelevant quadrant under the no cue condition; r_s (46) = -.35, $p = .01$. There was also a significant, positive correlation between visuospatial memory capacity and fixation percentage in the central interest area when all cue conditions combined; r_s (46) = .37, $p = .01$. Participants with better visuospatial memory tended to look more at the centre in all cue conditions and look less at the irrelevant quadrant in no cue condition.

4.6 Discussion

We investigated (1) whether looking at nothing increases as participants are asked to study and retrieve more difficult words in a yes/no recognition memory paradigm and if so, (2) which lexico-semantic variable(s) predict the change in memory-guided eye movements to relevant, blank location. We further tested how a visual cue presented between encoding and retrieval stages modulates the effects of word difficulty and word properties on memory performance and looking at nothing.

As shown in Experiment 1 (Kumcu & Thompson, 2016), participants displayed looking at nothing behaviour in congruent and no cue conditions but not in the incongruent cue condition. Incongruent cues functioned as interference. That is, the spatial index associated with the probe word was updated with a visual cue that did not match with the word's original location. In turn, the spatial index attached to the word and the spatial index attached to the visual cue competed and disrupted eye movements to blank locations. As looking at nothing behaviour was not exhibited in the incongruent cue condition, we investigated the effect of word difficulty and word properties on eye movements under congruent and no cue conditions.

We also examined memory performance under different retrieval conditions (cue conditions and high difficulty vs. low difficulty words) to verify the experimental manipulations and to compare against looking behaviour results. Unlike their effect on fixations, cue locations did not affect memory performance. Participants performed equally well under all retrieval conditions as demonstrated by both hit rate and hit latency. On the other hand, memory performance was superior in the retrieval of low difficulty words in comparison to high difficulty words as expected. Taken together, it is probable that visual cues were salient enough to modulate eye movements but not memory performance as memory performance was based more on the nonvisual, verbal parameters.

In line with our hypothesis, participants looked more at the blank, relevant quadrant compared to irrelevant quadrants when retrieving high difficulty words but not low difficulty words in congruent and no cue conditions. That is, participants relied on additional external sources when memory load was high, and they returned back to internal memory sources only when memory load decreased. We conceptualised memory load as the mental effort to maintain and reconstruct information as a function of difficulty. In this respect, such retrieval behaviour is in line with previous studies showing a proportional relation between memory load (via difficulty) and looking at nothing (Scholz et al., 2011; Wantz et al., 2015), where increase in load triggers eye movements. Additionally, the current study provides the first evidence that word difficulty as a function of lexico-semantic properties, modulates looking behaviour directly in memory.

Participants were likely to form stronger representations for low difficulty words. Retrieving these words required less mental effort as opposed to words that are more difficult to remember. Consequently, internal memory involving mental representations was sufficient to retrieve the probe word and solve the memory problem (i.e., yes/no judgement) accurately. However, integrated memory engaged spatial indices when retrieving difficult words. Thereby, the verbal component of the memory representations corresponding to high difficulty words was reinforced with stronger spatial information through eye movements and, as a consequence, memory load was alleviated (Ferreira et al., 2008).

High difficulty words in the current study were in fact more distinctive than low difficulty words in orthographic & phonological similarity, age of acquisition, availability and pronounceability. That is, high difficulty words were less similar with others in the lexicon, learnt at later ages, less available (i.e., do not come to mind easily) and less pronounceable. If high difficulty words were more distinctive in these variables, then why did participants still retrieve low difficulty words more accurately than high difficulty words? The variables which made high difficulty words more distinctive (i.e., orthographic & phonological similarity, age of acquisition, availability and pronounceability) did not play a role in hit rate in the current study. Major predictors of hit rate, and by extension we assume, memory load, were word length, imageability and concreteness. Low difficulty words were more imageable, more concrete and shorter (in letters and syllables). As a result, low difficulty words imposed less load on memory and were retrieved more accurately.

The results have important implications for theories postulating that cognitive work can be offloaded onto the environment (Risko & Gilbert, 2016) and for the dynamics of looking at nothing, in particular. First, we clearly showed that participants switched from internal to external sources swiftly and efficiently from trial to trial given that difficulty was a randomised, within-subjects variable in the current study. Such behaviour suggests that memory retrieval from internal and so-called "external memory" is not binary, but a dynamic process. Depending on the immediate memory load and strength of memory representations, retrieval behaviour hovers between the two extremes of a spectrum, where internal memory (representations) is on one end and external memory (spatial indices internalised via eye movements) on the other. Our results are in contrast to accounts that reject the existence of mental representations (Chemero, 2011), instead claiming that the external world is its own memory (e.g., O'Regan, 1992). Rather, findings support the position that the external world has a supportive role in memory (Ferreira et al., 2008; Johansson et al., 2011; Richardson et al., 2009). Eye movements are of particular importance in this pattern as they bind the internal representations to external information.

Whether or not eye movements are employed consciously as a memory strategy remains an unanswered question. There is evidence that explicit external support such as writing down to-be-remembered information is co-opted intentionally when it is offered as a choice (Risko & Dunn, 2015). We did not address this question in the current study. However, informal queries with the participants following the experiment revealed that they were not aware of the manipulation and did not look at the blank location with the intention of alleviating memory load (see also Johansson & Johansson, 2014; Scholz et al., 2014).

We did not present any evidence that looking at nothing has a functional role in memory. However, a number of studies demonstrated that looking at relevant, blank locations improves memory performance for both verbal and visual information (Johansson & Johansson, 2014; Scholz et al., 2018, 2016). What lies behind this discrepancy? We argue that the main reason is the difference in the experimental paradigms. Participants were instructed to look at either relevant or irrelevant quadrants during retrieval in the abovementioned studies evidencing functionality of looking at nothing. We did not design the current paradigm to test whether looking at nothing improves memory performance. Thus, eye movements were not manipulated, and all participants gazed freely during the retrieval phase. As a supplementary analysis, we tested whether fixations in the relevant, blank quadrant predict memory performance using mixed-effects models. As Martarelli et al. (2017) assert, the best way to understand functionality of looks to blank locations seems to be the manipulation of eye position at retrieval. In the current study, participants might have used other strategies to retrieve information from memory when they were not forced to look at certain positions on the screen.

Another possibility could be the difference in verbal information to be retrieved. In Scholz et al. (2017, 2014), participants encoded longer verbal information (i.e., factual sentences) and in the retrieval phase, a true/false statement probed participants' memory. In our study; however, participants were asked to encode four single nouns shown simultaneously and memory was probed with another single noun. In line with our findings showing a link between word difficulty and looking at nothing, we speculate that functional role of eye movements in memory might emerge when memory load reaches a certain threshold (see Johansson et al., 2012; Laeng et al., 2014; Spivey & Geng, 2000). It is important to highlight that although participants looked more at relevant, blank locations than irrelevant, blank locations when remembering difficult words, looking at nothing did not predict memory performance in the retrieval of difficult words either. It is possible that maintaining and retrieving single words instead of longer verbal information were demanding enough to elicit looking at nothing behaviour but not demanding enough to allow for functional eye movements. Given the previous evidence discussed above, it is highly probable that the role played by eye movements in memory is beyond an epiphenomenal by-product of the retrieval mechanism (see also Hannula et al., 2010). That said, future studies should test the functional role of eye movements in memory under different retrieval conditions and with different verbal or visual material in order to systematise the effect.

Lastly, we used the term “eye movements” to describe the looking behaviour in the present study, following the practice in the literature (e.g., Johansson & Johansson, 2014; Richardson et al., 2009; Spivey & Geng, 2000). That said, it should be noted that we did not present any evidence that eye movements, in the strictest sense, are relevant in looking at nothing for visually presented single words as opposed to looking at nothing during mental imagery (e.g., Brandt & Stark, 1997).

Which word properties contribute to looking at nothing? Within the scope of the second research question, we explored lexico-semantic variables predicting eye movements to blank locations by clustering the variables into three factors (imagery, lexical and length & similarity factors). Use of word properties as independent and continuous variables instead of within difficulty categories in mixed-effects models eliminated the possibility of any confounding influence of stimuli design on the results. Imagery, that is, imageability and concreteness was predictive of looking at blank locations during retrieval. Participants looked more at nothing when retrieving less imageable and more abstract words in no cue, that is, a “pure” looking at nothing condition. The effect of imageability and concreteness on fixations in the relevant quadrant under congruent cue condition was at a *p* level of .08. For the sake of simplicity, we will use the term “imageability” to refer both imageability and concreteness below.

Why did participants look more frequently to blank regions when retrieving less imageable words? As discussed in the introduction, imageability might have modulated eye movements in two different ways: due to its contribution to (1) word difficulty and thus, memory load or (2) mental imagery of words.

There is robust evidence that imageability is among the strongest predictors of performance in verbal recognition memory (Paivio, 1991). Accordingly, imageability might have affected fixations as the main moderator of word difficulty and memory load. However, retrieval performance was also predicted by length & similarity factor (length in letters, syllable length, phonological similarity and orthographic similarity) although it did not predict looking at nothing. This suggests that length & similarity (length, in particular) increased memory load as well. If the first account, that is, a difficulty/memory load account was indeed the only explanation for the effect of imageability on eye movements, length & similarity factor should have predicted eye movements as well.

As this was not the case, it appears more probable that imageability predicted looking at nothing mainly due to its contribution to the mental imagery of the words. We assume that participants relied more on mental imagery by looking at blank locations when the internal images activated by words fell short. This interpretation is also supported by the difference between congruent cue and no cue conditions in the effect of imageability on looking at nothing. The effect of imageability was revealed in the no cue condition but not in the congruent cue condition. In the same vein, the effect of mental imagery control emerged in the no cue condition but not in incongruent cue condition. As we discussed in the introduction, visual information emphasizing the location of the probe word could have interfered with the mental imagery process and minimised the effect of imageability on looking at nothing-

In a nutshell, weaker mental images corresponding to less imageable words were compensated for by looking more at nothing during retrieval. The effect of imageability on eye movements also disproves the prediction that participants treated the words as orthographic units when remembering them on a blank screen. It is safe to assume that participants formed and relied on conceptual mental images rather than images reflecting the physical properties of words. The role of mental imagery in the current study is noteworthy considering that the participants were not instructed to generate mental images to retrieve the words (cf., Laeng & Teodorescu, 2002). Such a retrieval behaviour supports grounded-embodied and perceptual approaches to memory suggesting that retrieval is, in essence, imagining and simulating the encoding (Albers et al., 2013; Glenberg, 1997; Jonides, Lacey, & Nee, 2005; Kent & Lamberts, 2008; Schacter et al., 2012; Wheeler et al., 2000).

4.7 Conclusion

From a wider perspective, reducing internal demands is a crucial and consistent function of behaviours in which cognitive work is externalised (Gilbert, 2015; Melinger & Kita, 2007; Schönpflug, 1986). Here, we analysed language-based factors which influence the propensity to engage the external world with eye movements to minimise memory load. Our findings show that people rely more on spatial indices when retrieving low imageable words. These findings can be considered compelling evidence for a flexible coordination of internal and external memory systems. Future studies should examine the temporal dynamics of the coordination between internal sources and external support.

4.8 Appendix

4.8.1 Outputs of best fit mixed-effects models for looking behaviour

Congruent cue

Table 4.3 Results of the mixed-effects model for fixation percentage in the relevant, blank quadrant under congruent cue condition offering the best fit by maximum likelihood

Fixed effects	β	SE (β)	*df*	*t*	*p*
(Intercept)	0.21	0.01	57.75	19.23	<.0001
Imagery	-0.02	0.01	138.23	1.76	.08
Random effects	Variance	SD			
Participants (intercept)	.002	0.04			
Imagery	< .001	0.01			
Length & similarity	< .001	0.02			
Items (intercept)	< .001	0.01			
Imagery	.004	0.06			
Residual	.085	0.29			
AIC	509.6				
BIC	584.9				

Model: *fixation percentage ~ imagery + (1 + imagery + length & similarity | participant) + (1 + imagery | item)*

No Cue

Table 4.4 Results of the mixed-effects model for fixation percentage in the relevant, blank quadrant under no cue condition offering the best fit by maximum likelihood

Fixed effects	β	SE (β)	*df*	*t*	*p*
(Intercept)	0.21	0.01	45.15	18.82	<.0001
Imagery	-0.02	0.01	296.05	2.03	.04
Random effects	Variance	SD			
Participants (intercept)	.002	0.05			
Imagery	< .001	0.00			
Length & similarity	< .001	0.01			
Items (intercept)	< .001	0.02			
Imagery	< .001	0.02			
Residual	.082	0.29			
AIC	473.4				
BIC	497.8				

Model: *fixation percentage ~ imagery + (1 + imagery + length & similarity | participant) + (1 + imagery | item)*

4.8.2 Outputs of best fit mixed-effects models for memory performance

Hit rate

Table 4.5 Results of the mixed-effects model for hit rate offering the best fit by maximum likelihood

Fixed effects	β	SE (β)	z	p
(Intercept)	1.46	0.11	13.57	<.0001
Length & similarity	0.24	0.05	4.86	<.0001
Imagery	0.18	0.04	3.96	<.0001
Random effects	Variance	SD		
Participants (intercept)	0.46	0.68		
Imagery	0.01	0.11		
Lexical	0.08	0.27		
Length & similarity	0.02	0.13		
Items (intercept)	0.003	0.06		
Imagery	0.002	0.05		
Lexical	0.01	0.11		
Length & similarity	0.01	0.12		
AIC	4,344.1			
BIC	4,490.6			

Model*: hit rate ~ imagery + length & similarity + (1 + imagery + lexical + length & similarity | participant) + (1 + imagery + lexical + length & similarity | item)*

Hit latency

Table 4.6 Results of the mixed-effects model for hit latency offering the best fit by maximum likelihood

Fixed effects	*β*	SE (*β*)	*df*	*t*	*p*
(Intercept)	2044.16	68.16	48.02	19.24	<.0001
Length & similarity	-83.24	17.58	322.78	4.73	<.0001
Imagery	-48.66	20.41	44.89	2.38	.02
Lexical	41.49	19.51	102.85	2.13	.04
Random effects	Variance	SD			
Participants (intercept)	208,523.2	456.64			
Imagery	5448.7	73.82			
Lexical	4291.7	65.51			
Length & similarity	778.5	27.90			
Items (intercept)	1393.9	37.33			
Imagery	2132.4	46.18			
Length & similarity	527.0	22.96			
Residual	971,565.0	985.68			
AIC	56,413				
BIC	56,541.6				

Model*: hit rate ~ imagery + length & similarity + (1 + imagery + lexical + length & similarity | participant) + (1 + imagery + length & similarity | item)*

Chapter 5

Experiment 3

Simulating Space with Language: Horizontal and Vertical Coordinates for 1439 English Nouns

5.1 Motivation and Aims

Simulation of word locations in the absence of words were explored in Chapter 3 and 4 (Experiment 1 and 2). In the following chapters (Experiment 3 and Experiment 4), we shift our focus to spatial simulations evoked by words directly upon their presentation (see Chapter 2.1.3). The study reported in this chapter is a norming study that has two aims: (1) Exploring the consistency of spatial mappings between words and the simulated locations they trigger in a systematic manner. (2) collecting spatio-lexical norms for controlling the words to be used in Experiment 4 and future studies.

5.2 Abstract

The link between language and space is well-recognised (Lakoff & Johnson, 1980b; Talmy, 1983). There is now mounting evidence that words are associated with locations in space that play a role in visual and semantic processing (e.g., Meyer & Robinson, 2004). Despite the wealth of theoretical and experimental evidence, only a few studies have provided spatial norming for words (e.g., Meteyard & Vigliocco, 2009). Here, we present horizontal and vertical location norms for 1439 concrete and abstract English nouns. Participants were asked to give their location ratings on a two-dimensional coordinate system. Results show that the majority of the words in our pool were located on a diagonal line from bottom left to top right and participants were more conservative when using the horizontal plane as opposed to vertical plane. Both horizontal and vertical positions were positively correlated with goodness. We observed a remarkable consistency between participants in ratings. Findings are discussed

within the embodied perspectives of language framework related to spatial mapping. Spatial norms can be viewed and/or downloaded at https://osf.io/wb6pm/.

Highlights

- Both concrete and abstract nouns were consistently associated with horizontal and vertical positions on a two-dimensional coordinate system.
- Participants tended to associate "good words" with higher and rightward positions.
- Vertical space was more dominant than horizontal plane in spatial mappings.

5.3 Introduction

Many concepts have spatial associations. Expressions in everyday language such as "prices have recently gone *up*", "she felt *deep* sadness", "holiday is fast *approaching*" are clear examples that the spatial arrangement of language is deeply entrenched in cognition. Linguistic units corresponding to object and concepts, words and nouns, in particular, exhibit highly visible associations with space (e.g., roof - up).

Mappings between words and space have been frequently explained within embodied perspectives. Perceptual (e.g., Pulvermüller, 1999) and developmental (e.g., Piaget & Inhelder, 1969) theories of lexical representation posit that humans form sensorimotor experiences before the acquisition of language by tasting, touching, hearing, smelling or positioning the objects around them. As language is acquired and abstract thought is developed in time, verbal representations are coupled with prelinguistic sensorimotor memories. As a consequence, words can activate perceptual simulations (Barsalou, 1999, 2008; Gallese & Lakoff, 2005; Pecher, Boot, & Van Dantzig, 2011; Pecher & Zwaan, 2010). For example, reading the word "apple" activates previously acquired sensory experiences associated with the apple as a fruit;

such as taste, smell, feel, colour and shape etc. Likewise, many common words can trigger spatial simulations and thus, are associated with locations in space. For instance, the word, "sun" is associated with upward locations as the actual location of the sun in the physical world (in addition to other sensory experiences associated with it such as heat) is rooted in the representational system and tightly coupled with the verbal representation of the word. Alternatively, words such as "worm", "foot", "lake" tend to be associated with downward locations.

In addition to object words, pairings between abstract words and space manifest in language as well. Sensorimotor representations are more concrete and thus, more efficient in expressing abstract concepts. *Conceptual metaphor theory* (Gibbs, 1994; Lakoff & Johnson, 1980a, 1999; see also Landau, Meier, & Keefer, 2010) suggests that people tend to ground abstract concepts in concrete concepts through metaphors including metaphors involving space and orientation. As one of the most salient examples, goodness and the words with positive valence (e.g., honesty) are associated with upward locations, while negative words (e.g., murder) are mapped onto downward locations (i.e., good is up and bad is down). Such a mapping occurs as physical highness provides a concrete schema for abstract concepts likely due to our physical positioning in space, a sensual experience we have had since birth.

There is a wealth of experimental evidence that abstract words with valance are mapped onto vertical space, which in turn, has an effect on visual and semantic processing. For example, reading positive or negative words as primes orients attention to higher or lower locations in space influencing decisions on subsequent words even if reading emotion-laden words is not related with the main task. Specifically, people are faster to detect higher targets followed by a positive word (*conceptual cueing effect*) (Gozli, Chasteen, & Pratt, 2012; Gozli, Pratt, Martin, & Chasteen, 2016; Zanolie et al., 2012) and evaluate valence words faster if they are presented in congruent locations (Meyer & Robinson, 2004). More specifically; concepts of power (Schubert, 2005; Zanolie et al., 2012), divinity (Chasteen et al., 2010), healing (Leitan et al., 2015) and self-esteem (J. E. T. Taylor et al., 2015) are described as a vertical dimension in space in that words denoting powerfulness (e.g., boss, judge, professor etc.), God (e.g., Lord, almighty etc.), higher self-esteem (e.g., confident, proud, assertive etc.) and health (rejuvenated, alive, fit etc.) are associated with higher locations (and vice-versa) with similar cueing effects.

Concepts are represented on the horizontal dimension as well. For instance, temporal words (e.g., yesterday, next, previously etc.) are projected onto the horizontal axis, where past is left, and future is right (Boroditsky, 2000; Casasanto & Boroditsky, 2008; Santiago, Lupiáñez, Pérez, & Funes, 2007; Weger & Pratt, 2008). Numbers (e.g., 6) and number words (e.g., six) are internally represented on a mental number line from left to right according to their magnitude (i.e., lower digits = leftward positions and higher digits = rightward positions), which in turn shifts attention to relative parts of the external space (Martin H Fischer, Castel, Dodd, & Pratt, 2003). The way people link space with language seems to be related to their bodies (*body specifity hypothesis*). For example, people allocate positive concepts on the dominant hand-side of their bodily space (i.e., right = good for right-handers and left = good for left handers) (Casasanto, 2009; de la Vega, de Filippis, Lachmair, Dudschig, & Kaup, 2012; De la Vega, Dudschig, De Filippis, Lachmair, & Kaup, 2013).

Although the link between space and language is well-established on both experimental and theoretical grounds, only a few studies provided location norms for words either as a standalone norming study or as a pre-test to develop stimuli. In the only standalone norming study to our knowledge, Meteyard and Vigliocco (2009) asked 96 native English speakers to read rebus sentences where geometric shapes were used instead of subjects and objects. Sentences involved transitive verbs such as "[circle] *writes* [square]" or intransitive verbs such as "[circle] *advances*". Then, participants were shown diagrams, which represented possible motions and directions (toward/away from the subject) and positioning (top/bottom/right/left) arrangements with the circle and square. Participants chose the most appropriate diagrams describing the verb (see also Richardson et al., 2001). Thereby, axis and directions norms according to the diagram preference ranks were provided for 299 English verbs. Other studies frequently referred to spatial norming as a pre-test to manipulate and/or control spatial iconicity of the verbal material to be used in the actual experiments and thus, norms were not generally presented as supplementary: For instance, Marmolejo-Ramos et al. (2013) instructed participants to place 64 personality-trait adjectives in English and Japanese on a two by two grid through a task with a cover story (Experiment 2), in which they chose the best candidate for a job by allocating the adjectives to candidates on the grid. Thus, x and y coordinates were computed in visual angles of 64 adjectives on the computer screen. A number of studies relied on Likert-scales in spatial norming pre-tests: In Lachmair, Dudschig, De Filippis, de la Vega and Kaup (2011), 49 volunteers rated 104 German nouns with respect to the referents' typical location, using a 5-point Likert-scale ranging from down to up. Dunn, Kamide and Scheepers (2014) instructed Internet-based participants to rate 402 English words for vertical associations on a 11-point bipolar scale ranging from - 5 (labelled as down) to + 5 (labelled as up), where the midpoint (0) was labelled as neutral (no vertical association). In a similar fashion, 100 Internet-based participants rated 180 concrete nouns in English on the vertical scale on a scale from 1 (extremely low) to 5 (extremely high) (pre-test of Experiment 2) in Estes, Verges and Adelman (2015). Although not a location norming study in its strictest sense, Louwerse (2008) asked the participants to rate the spatial iconicity of word pairs (e.g., bridge - river); more specifically, the likelihood that one word appeared above the other on a scale of 1-6, with 1 being extremely unlikely and 6 being extremely likely.

In order to fill the gaps in the previous spatial norming studies discussed above, we aimed to provide precise location ratings on both horizontal and vertical dimensions for a larger set of English nouns involving both abstract and concrete nouns. Our study differs from the previous studies on two main accounts: **(1) Two-dimensional coordinate:** Rather than using a Likert-scale with 5, 6 or 11 points either on vertical or horizontal direction, we provided the participants a Cartesian coordinate system, whose horizontal and vertical axes were 960 pixels long (+/- 460 pixels on both directions from the origin), giving 921600 different clickable points to locate the words (see Figure 5.1). **(2) Ratings on both horizontal and vertical dimensions:** As a result of using a coordinate system, location ratings for each word were provided as a (x, y) coordinate. For example, if the word "sun" is at (31, 370), it means it is located on the rightward side of the horizontal scale (x) as of 31 pixels and upward side of the vertical scale (y) as of 370 pixels. Likewise, a word at (-230, -12) means that the word is located on the leftward side of the horizontal scale (x) as of 230 pixels and downward side of the vertical scale (y) as of 12 pixels. Further, participants were directly instructed to specify the locations of words in our study without any additional tasks. Thereby, we aimed to access relatively conscious representations of verbal - spatial mappings.

5.4 Method

5.4.1 Participants

A group of thirty undergraduate psychology students at the University of Birmingham (five males; M_{age} = 19, SD = 0.69, range: 18 - 20) took part for course credit. All participants were monolingual native speakers of British English (speaking/learning only English from birth and currently using English as their primary language) as determined with the Language History Questionnaire (version 2.0; Li, Zhang, Tsai, & Puls, 2013). Thus, participants can be regarded as Western European culturally.

Handedness was tested with Edinburgh Handedness Inventory. We did not address handedness as an additional variable and control the participants accordingly, but we aimed to observe the variability among participants. Results showed that there were 16 right-handed, one left-handed and 13 ambidextrous participants (i.e., using both right and left hands) in the study. Mean laterality quotient was 65 on a scale of -100 (absolute left-handedness) to +100 (absolute right-handedness) showing that right-handedness was dominant among participants in particular including participants who can use their both hands, which is a frequently seen case considering the prevalence of right-handedness (Corballis, 2003).

Participants reported normal or corrected-to-normal vision, no speech or hearing difficulties and no history of any neurological disorder. All participants were fully informed about the details of the experimental procedure and gave written consent. Post-experiment debriefing revealed that all participants were naïve to the purpose of the experiment.

5.4.2 Materials and apparatus

1482 nouns (excluding 2 pseudowords) were included in the study. Words were compiled from three main sources: (1) extensions of Paivio, Yuille and Madigan norms for 925 nouns (J. M. Clark & Paivio, 2004), (2) modal exclusivity norms for 400 nouns (Lynott & Connell, 2013), (3) material in two spatial iconicity studies (Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013; Zwaan & Yaxley, 2003). 151 nouns, which are not available in these sources, were selected for their potential emotional (e.g., “despair”) or spatial content (e.g., “north”).

Stimuli were presented on a TFT LCD 24-inch widescreen monitor operating at 144 Hz with a resolution of 1920 x 1080 pixels (508 mm x 285.75 mm). The experiment was programmed in and run on OpenSesame 3.1.2 (Mathôt, Schreij, & Theeuwes, 2012). Data were analysed and visualised with R (version 3.3.2) (R Core Team, 2017).

5.4.3 Procedure

The paradigm was modelled after the norming study of Brysbaert, Warriner and Kuperman (2014). Material set was randomly distributed over 5 lists of 290 words. We aimed to compute internal consistency of the data. Thus, each list additionally included 32 control words, which were rated by all participants. Further, two pseudo words ("reiltas" and "vasagle") were added to the lists to distinguish and exclude participants who were not attending the task. In total, each participant rated 324 words. Control words were rated by 30 participants and different words in the lists were rated by six participants.

A list of 10 practice words was presented to participants before each set. The practice words represented the entire spatial range (i.e., top left, top right, bottom left, bottom right, left, right, up, down, neutral) with the aim of introducing the participants to the variety of the words they would encounter. The control words were also from the whole range and used to test consistency between participants and throughout the experiment. They were scattered randomly throughout the lists. The experiment was composed of two consecutive phases (see Figure 5.1 for a schematic illustration of the experimental design, see Appendix (5.8.1) for the instructions of the experiment): **(1) Study:** Participants were presented a centrally placed single word (Courier, font size = 52) in capital letters. They were instructed to read the word carefully and visualise its meaning before moving onto the rating phase. There was no time limit for the study phase and it ended by hitting the space key. **(2) Rating:** Participants were presented a rating screen during the rating phase. The rating screen included a blank, central square of 480 x 480 in pixels (11° x 11° of visual angle). Four direction labels (i.e., up, down, left and right) were placed outside the rating square. Participants were expected to identify the word's location by left clicking any point within the square. When a participant clicked a location, the experimental software automatically recorded click location in horizontal and vertical (x, y) coordinate space. A smaller box (32 x 32 in pixels) outside the rating square was placed at the bottom right of the rating screen. Participants were instructed to click this box when they do not know the word enough to give a rating.

The order of words in the lists was fully randomised. A typical session lasted approximately 30 minutes. All participants labelled pseudo words as unknown, suggesting a good level of attention.

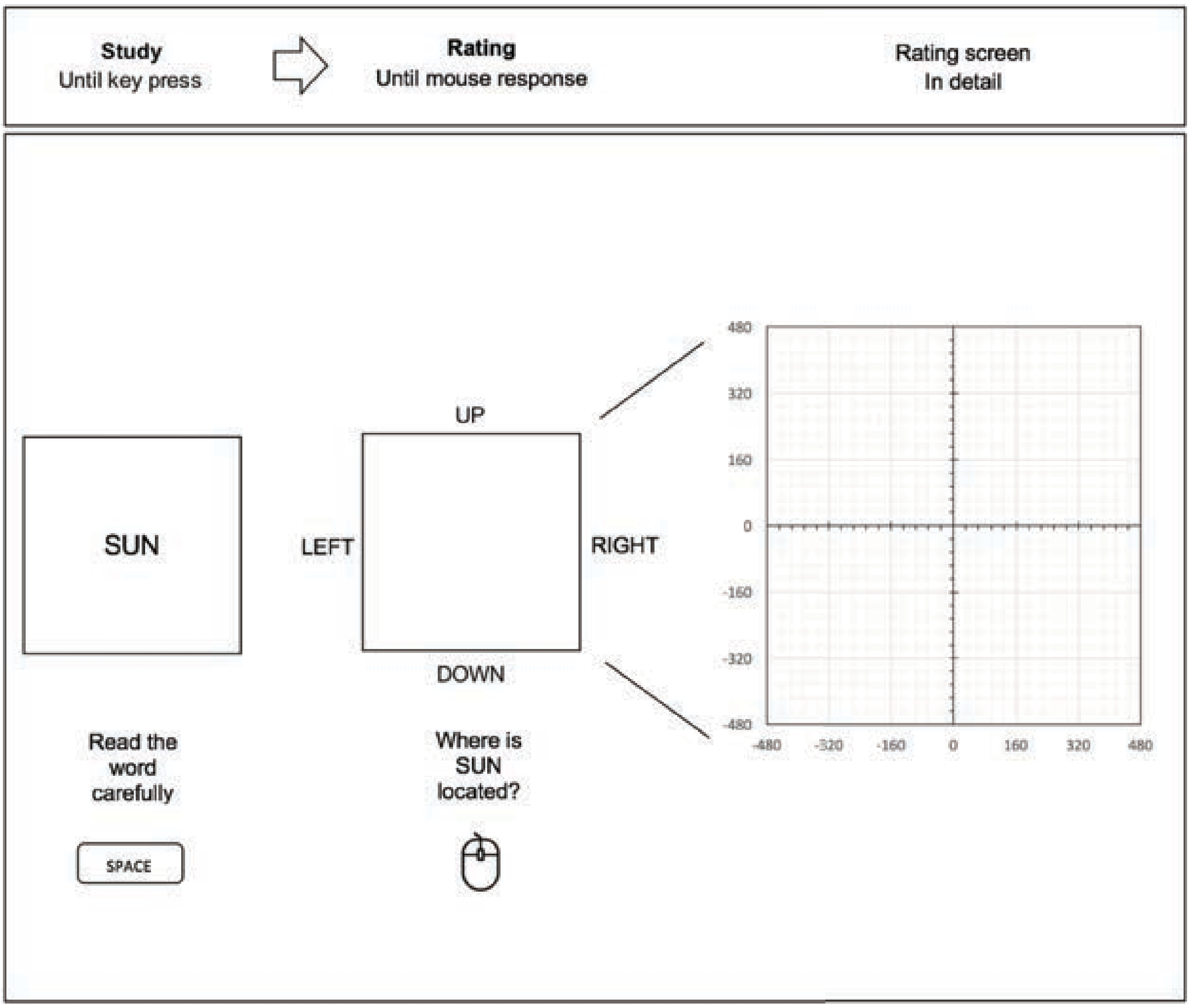

Figure 5.1 Diagram showing the experimental procedure and the coordinate system behind the rating box. Precision of the coordinate system was downscaled for visibility. The coordinates were not visible during rating and participants saw a blank screen.

5.5 Results

5.5.1 Data trimming

9925 ratings were collected from 31 participants. One participant was removed due to missing responses and one participant was replaced due to high percentage of words rated as unknown (78%). Overall, 9720 unique ratings from 30 participants were analysed. 43 nouns were excluded as they were rated as unknown by four or more participants in each list (see Appendix (5.8.3)). As a result, 1439 nouns were given a location rating. Horizontal and vertical ratings for each word were included in the analyses.

5.5.2 Reliability and validity

We calculated intraclass correlation coefficients (ICC) via a two-way random absolute agreement model. Participants rating the control words showed excellent interrater reliability for horizontal (ICC = 0.95, *CI* (95%) = 0.92 - 0.97) and vertical positions (ICC = 0.98, *CI* (95%) = 0.97 - 0.99) as determined by the cut-offs provided by Cicchetti (1994) and Hallgren (2012).

We grouped control words into two groups as concrete and abstract words based on concreteness (range: 2.87 – 6.77). Participants rating the control words showed excellent interrater reliability for horizontal positions of concrete words (ICC = 0.96, *CI* (95%) = 0.92 - 0.98) and for horizontal positions of abstract words (ICC = 0.94, *CI* (95%) = 0.89 - 0.98). Participants also showed excellent interrater reliability for vertical positions of concrete words (ICC = 0.98, *CI* (95%) = 0.96 - 0.99) and for vertical positions of abstract words (ICC = 0.98, *CI* (95%) = 0.96 - 0.99). Point estimate of ICCs for concrete/abstract words and CIs of ICCs for concrete/abstract words overlapped in each case, suggesting that there was not a significant difference between the consistency among participants when rating concrete or abstract words on the horizontal or vertical axis.

There was a high degree of consistency between participants in location ratings. The mean standard error for all nouns was 3.02 for horizontal plane and 5.37 for vertical plane. We also computed rater correlation coefficients between each participant's average rating for a given word and the mean rating for the word in question within each presentation list. The mean coefficient of the correlation sets was 0.76 (SD = 0.07, range: 0.59 - 0.84) for the horizontal position and 0.65 (SD = 0.12, range: 0.36 - 0.82) for the vertical position.

Due to the limited number of location rating studies, we could only verify the external validity of our vertical ratings by correlating them with a data set from Estes et al. (2015). Overlapping 94 words showed a strong, positive correlation; r (92) = .92, p < .0001 (see Figure 5.2). Experimental validation of the norms is provided in Experiment 4, in the next chapter.

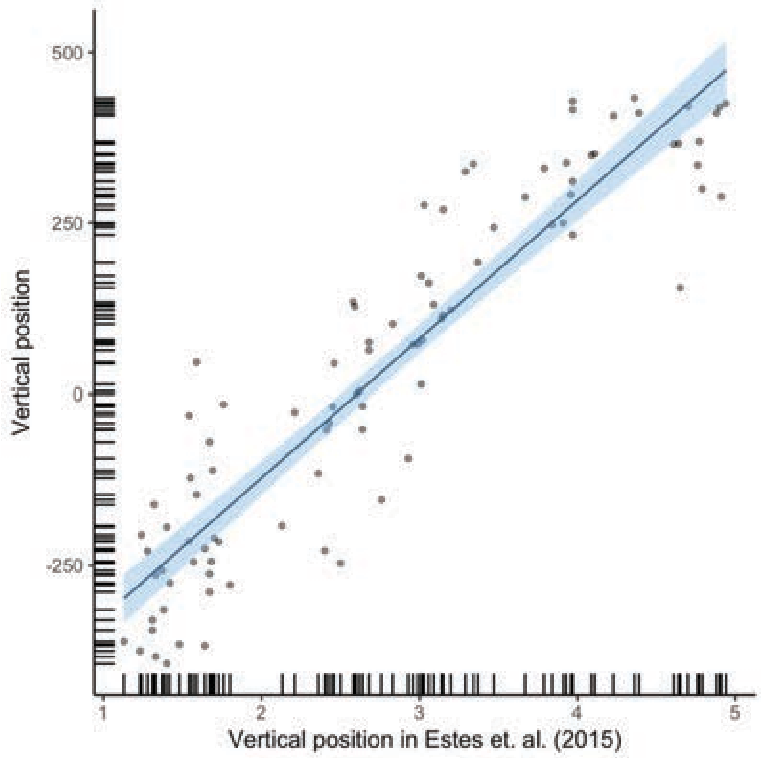

Figure 5.2 Correlation between vertical position ratings in the data set and Estes et al. (2015). Scatterplot has a linear regression line. Blue band around the line represents 95% confidence interval. Tassels at the x and y axis illustrate the marginal distribution of data along vertical position ratings in Estes et al. (2015) and vertical position ratings in the current study.

5.5.3 Descriptive statistics

Descriptive statistics of the ratings for horizontal and vertical positions were presented in Table 5.1. Analyses of symmetry revealed relatively bell-shaped symmetrical curves for both horizontal and vertical positions. Along with that, both horizontal and vertical positions indicated slight tendencies for left skewness, where the means were located on the left side of the medians. Density plots (see Figure 5.3 - A), variance and symmetry measures clearly showed that ratings on the vertical axis were more spread out than the ratings on the horizontal axis. In other words, participants tended use the whole range of the vertical scale but rated the words more conservatively on the horizontal scale. Heatmap showing the distribution of the words across the coordinate system indicated that the majority of the words were rated on a diagonal axis from bottom left to top right (see Figure 5.3 - B). This was also evident from the strong, positive correlation between horizontal and vertical positions; r (1437) = .44, $p < .0001$. Words of particular interest are presented in Figure 5.4 and Appendix (5.8.2).

Table 5.1 Descriptive statistics of 1439 words

Statistic	Horizontal position	Vertical position
Mean	6.14	21.64
SD	114.58	203.81
SE	3.02	5.37
Median	11.17	30.75
Trimmed mean	8.69	25.73
Median absolute deviation	101.31	208.93
Minimum	-471.67	-458.50
Maximum	453.83	452.80
Range	925.50	911.3
Skewness	-0.24	-0.16
Variance	13128.40	41536.82
Kurtosis	1.03	-0.64
CI of mean (95%)	0.21 - 12.06	11.10 - 32.18

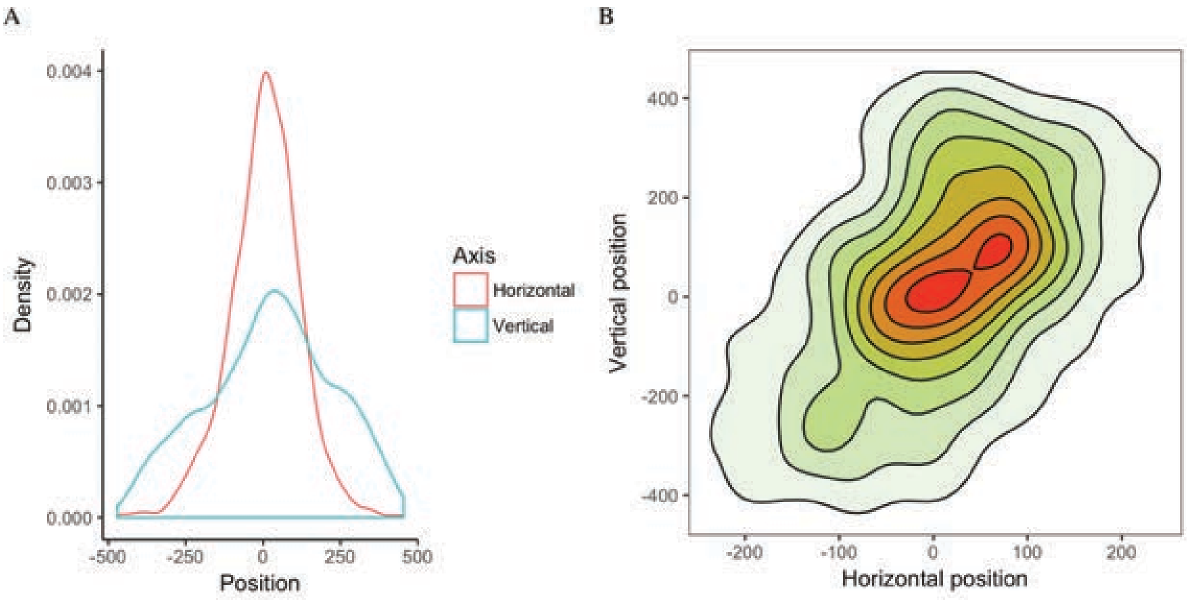

Figure 5.3 A) Density plot and B) heatmap of the word locations on the horizontal and vertical axis. Density plot shows the distribution of ratings over the positions on the horizontal and vertical axis. Heat map shows the same distribution but as two-dimensional and on the rating screen. Colours represent the frequency of the positions where participants clicked to rate the words. Accordingly, lighter colours indicate low frequency and darker colours indicate high frequency. Note that blank area outside the coloured pattern does not refer to absence of ratings but rating frequency under a certain threshold.

5.5.4 Spatiality and controversiality

We computed a variable, *spatiality*, for each word based on the coordinates. Spatiality shows the distance of a word from the centre. Hence, it is the extent to which a word is spatial as opposed to being spatially neutral. The spatiality value is the sum of the horizontal and vertical locations ratings (negative values were transformed into positive), normalised into a range from 0 to 1. Thus, higher values indicate higher spatiality (see Table 5.3). Another variable, *controversiality*, was computed to investigate the words that participants rated on vastly different locations of the coordinate system. The controversiality value is the average of the standard deviations between ratings among participants within each list, normalised into a range from 0 to 1. Accordingly, higher values indicate higher controversiality in ratings (see

Table 5.3). As opposed to standard error, it is based on the participants, not the words in the list. Descriptive statistics for both variables are presented in Table 5.2.

Table 5.2 Descriptive statistics of spatiality and controversiality (all measures range between 1 and 0)

Statistic	Spatiality	Contraversiality
Mean	0.33	0.37
SD	0.19	0.12
Median	0.31	0.37
Trimmed mean	8.69	0.37
Median absolute deviation	0.21	0.12
Skewness	0.52	0.28
Kurtosis	-0.30	0.77
SE	0.01	0.00
Variance	0.04	0.02
CI of mean (95%)	0.32 - 0.34	0.37 - 0.40

Table 5.3 Words sorted according to spatiality and controversiality (all measures range between 1 and 0)

	Word	Spatiality	Word	Controversiality
Highest	Northwest	1	Arbiter	1
	Murder	0.95	Corner	0.79
	Suppression	0.93	Border	0.78
	Justice	0.92	Catacomb	0.77
	Grief	0.91	Close	0.76
Lowest	Centre	0	Left	0
	Site	0.005	Centre	0.0009
	Folly	0.006	Right	0.015
	Odour	0.007	West	0.019
	Pair	0.008	Down	0.02

Figure 5.4 Scatterplot showing the word locations on the horizontal and vertical axes. Dot colours represent the controversiality of the words where blue indicates the most consistent and red indicates the most controversial words. An interactive version of the scatterplot can be viewed at https://plot.ly/~alperkumcu/8/.

5.5.5 Relationship with other variables

Correlations between the location ratings and lexico-semantic variables (J. M. Clark & Paivio, 2004) were examined with the 883 nouns that occurred in both studies. There was a strong, positive correlation between goodness of a word and its horizontal position; r (1437) = .50, $p < .0001$ and vertical position; r (1437) = .64, $p < .0001$. In other words, positive words were associated with upward and rightward locations; whereas, negative words were associated with downward and leftward locations. There was also a weak but positive correlation between frequency and vertical position; r (1437) = .21, $p < .0001$. Upward words tended to be more frequent than downward words. Words with higher spatiality were more emotional; r (1437) = .37, $p < .0001$ and more abstract; r (1437) = -.24, $p < .0001$. Other correlations were either weak or negligible. Correlations are shown in Figure 5.5.

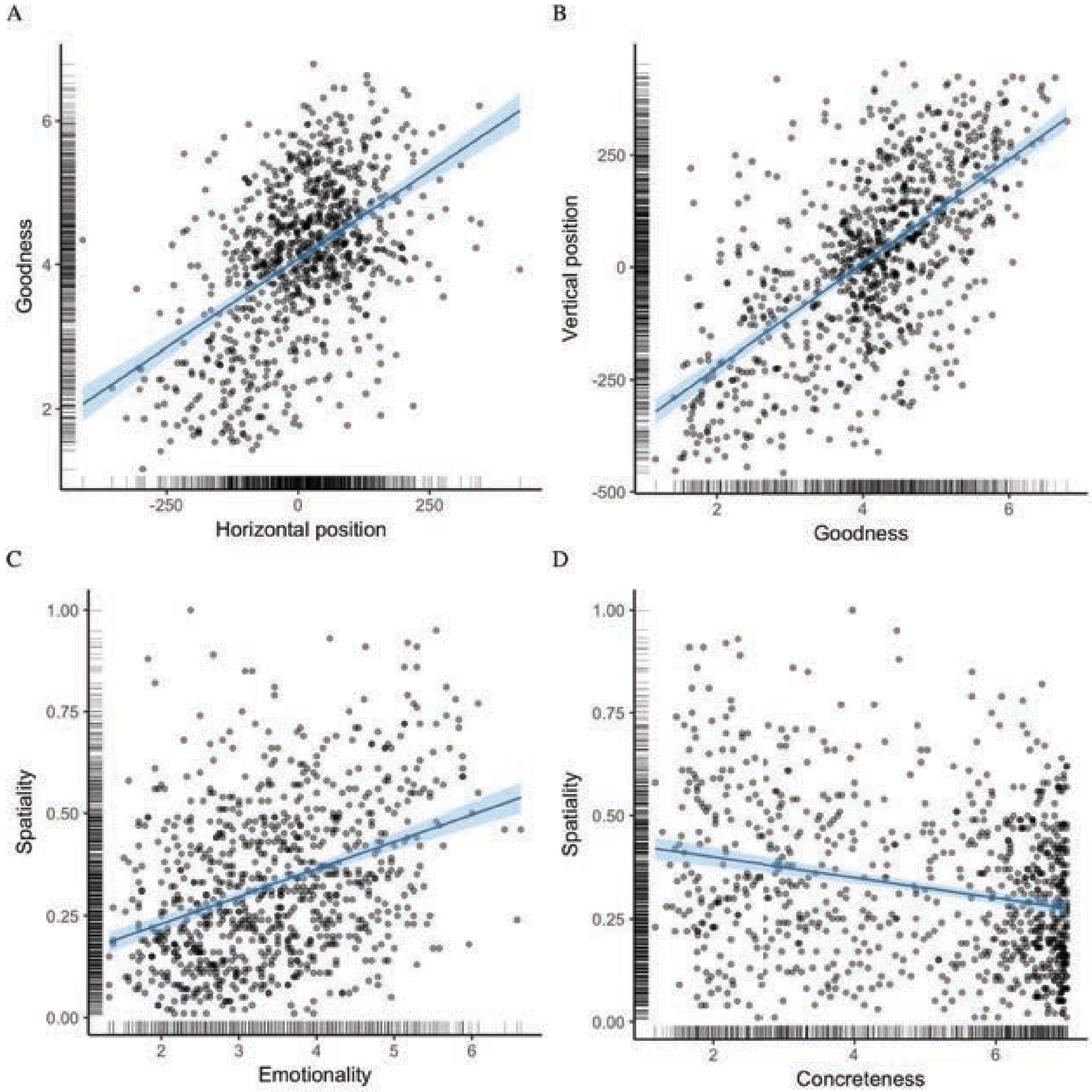

Figure 5.5 Scatterplots showing the correlations between (A) goodness and horizontal position (B) goodness and vertical position (C) spatiality and emotionality (D) spatiality and concreteness. Scatterplots have linear regression lines. Blue bands around the lines represents 95% confidence interval. Tassels at the x and y axis illustrate the marginal distribution of data along x and y variables.

5.5.6 Spatial labels

We divided the horizontal and vertical axes into three equal pieces (each piece is 160 pixels long) and tagged the words with 12 different labels as a combination of five zones (i.e.; up, down, left, right and centre) and three degrees of spatiality (extreme and medium and only for the central zone, low) (see Figure 5.6 and Table 5.4). All words are located in the combination of two zones as they have both horizontal and vertical ratings. For example, “sun” is located at (30.53, 369.5); which means, it is a top, neutral rightward word located in the upper central zone. Spatial measures of words grouped into zones were presented in Table 5.5.

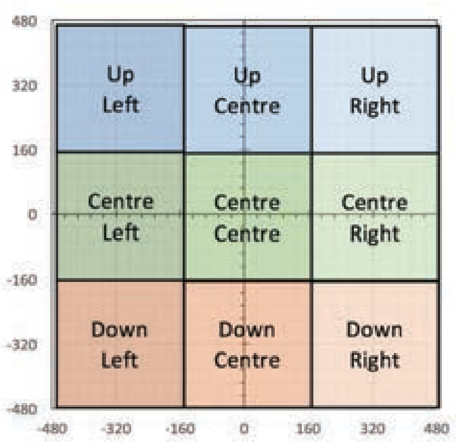

Figure 5.6 Zones on the coordinate system.

Table 5.4 Spatial labels and their spatial ranges

Zone	Degree	Label	Range (in pixels)	Dimension
Up	Extreme	Top	[480, 321]	Vertical
	Medium	Higher	[320, 161]	Vertical
Down	Extreme	Bottom	[-480, -321]	Vertical
	Medium	Lower	[-320, -161]	Vertical
Left	Extreme	Far left	[-480, -321]	Horizontal
	Medium	Leftward	[-320, -161]	Horizontal
Right	Extreme	Far right	[480, 321]	Horizontal
	Medium	Rightward	[320, 161]	Horizontal
Centre	Low	Neutral higher	[160, 1]	Vertical
	Low	Neutral lower	[-160, -1]	Vertical
	Low	Neutral leftward	[-160, -1]	Horizontal
	Low	Neutral rightward	[160, 1]	Horizontal
	Low	Neutral*	0	Horizontal and vertical

Ranges are inclusive of the thresholds. Words with location ratings between the thresholds (e.g., 320.38, -161.60 etc.) were rounded above or below and labelled accordingly.

* Words with a location rating of 0 on either horizontal (n = 7) and vertical plane (n = 6) were labelled as “neutral” without any spatial label.

Table 5.5 Means (standard errors) of spatial measures with regard to words in zones

Zone	Number of words	Horizontal position	Vertical position	Spatiality	Contraversiality
Up Left	4	-272.81 (54.64)	289.54 (30.81)	0.74 (0.11)	0.44 (0.13)
Up Centre	307	39.94 (3.58)	277.28 (4.24)	0.44 (0.01)	0.34 (0.01)
Up Right	58	216.15 (6.15)	274.35 (9.33)	0.64 (0.01)	0.36 (0.01)
Centre Left	46	-235.06 (11.87)	-40.12 (12.2)	0.41 (0.02)	0.41 (0.02)
Centre Centre	694	12.62 (2.81)	15.1 (3.22)	0.18 (0)	0.39 (0)
Centre Right	39	227.64 (10.58)	37.63 (12.42)	0.39 (0.01)	0.42 (0.02)
Down Left	60	-222.93 (5.58)	-308.78 (10.2)	0.7 (0.02)	0.36 (0.01)
Down Centre	226	-42.63 (4.91)	-272.28 (4.87)	0.45 (0.01)	0.37 (0.01)
Down Right	5	263.03 (51.48)	-220.13 (18.57)	0.63 (0.07)	0.54 (0.12)

5.6 Discussion

We presented location norms on the horizontal and vertical axes for 1439 English nouns. The main findings are summarised and interpreted below.

Words were rated on a diagonal line from bottom left to top right in our data set. Further, variance on the vertical space was considerably higher. As a result, top right and bottom left zones emerged as conventional zones as opposed to top left and bottom right zones. Thus, we conclude that participants relied more on the vertical space when associating the words with the locations. The dominance of the vertical plane over the horizontal when both dimensions are investigated separately was evidenced before (Franklin & Tversky, 1990; Marmolejo-Ramos et al., 2013). For instance, the “up is good” metaphor is more salient than the “right is up” metaphor as shown with emotional face targets (Damjanovic & Santiago, 2016). Here, we demonstrate that people project words on the vertical space more frequently than horizontal space even when two axes are available at the same time. The distribution of words on the coordinate system around the diagonal axis can be explained with the relationship between goodness and spatial positions. As expected, goodness was positively correlated not only with vertical but also with horizontal positions. As words got more positive in valence, they were more likely to be located on the top right position and as they got "worse", they were more likely to be located on the bottom left. Thus, participants utilised the horizontal space when grounding goodness as well; yet, not as liberally as the vertical space. A strong, positive correlation between goodness and horizontal space can be attributed to the body-specifity hypothesis (Casasanto, 2009). Almost all participants except for one were either right-handed or ambidextrous in our study. As a result, positive words were mapped onto the right side of the coordinate system. We did not address the handedness as a variable and did not balance the raters accordingly in the current study. Further research can make use of the coordinate paradigm to investigate the impact of handedness on the representation of valence words on the horizontal axis.

In addition to abstract words with valence; object words, whose referents occur in particular locations in the physical world, were rated on their encoded locations and particularly on the vertical space (e.g., tower, foot, hair, cellar etc.). On the other hand, horizontal space was mostly occupied by the nouns denoting time (e.g., past, yesterday, history etc.), direction (e.g., route) and motion (e.g., car). There is compelling evidence that the temporal words are mapped onto horizontal space (e.g., Weger & Pratt, 2008) yet not for direction and motion concepts. Along with that, such a tendency is not non-intuitive given that diagrams with horizontal arrows represent mental concepts such as order, cause, motion and outcome (Tversky, 2009).

Our results show the spatial words (i.e., words away from the centre) were also more emotional and abstract. These findings are expected considering how emotion, regardless of being positive or negative, is transferred to space and that emotional words are typically abstract (e.g., Meyer & Robinson, 2004). Yet, it is worth considering that many models of affect represent arousal, as a dimension of emotion, on the vertical space as well (e.g., annoyance vs. rage) (Rubin & Talarico, 2009; Yik, Russell, & Barrett, 1999). This might account for why emotionality increases as we move towards to the edges and borders of the coordinate system. We assume that participants rated the words higher/lower or more leftward/rightward as arousal level associated with the words increased.

Diversity of word categories is an important advantage of the present study. Participants rated concrete words referring to objects perceived in typical locations in physical world such as “bird”, abstract words with metaphorical spatial associations such as “honesty” and also, emotionally neutral words which explicitly shows positions in the subject’s immediate environment such as “east”, “left” or “bottom”. In this respect, our ratings suggest that mapping associations between language and space are highly prevalent in that they spread over different types of words. We argue that word spatiality might serve as an important variable to manipulate or control in future studies.

It is important to note that spatiality and spatial iconicity measure highly similar but different concepts. Spatial iconicity refers to the strength of the mapping between the word and its location (e.g., Louwerse, 2008); whereas, spatiality shows the extent to which a word is distant from being spatially neutral. Thus, words with high spatiality might not necessarily have high spatial iconicity. For example, “northwest” and “murder” are the most spatial words in our study although intuitively, “northwest” seems to have a higher spatial iconicity than “murder”. Future research should investigate the relationship between spatiality and spatial iconicity.

The most prominent finding is the reasonably high degree of agreement between participants in location ratings. Humans do not learn locations when they learn words. Although object words are mapped onto space implicitly, there are not explicit grammatical or lexical conventions in language with regard to associations between space and language. Even though, participants systematically and consistently located the words on certain points regardless of spatiality. Language-driven perspectives of language - space correspondence (Goodhew, McGaw, & Kidd, 2014; Louwerse, 2008; Zwaan & Yaxley, 2003) as an alternative to an embodied account attributes the link between language and space to language use; more precisely, the frequency with which abstract and spatial words co-occur in language. Accordingly, certain abstract concepts are associated with locations due to the frequent use of spatial terms with these abstract concepts (e.g., God *above*, feeling *down*, look *forward*). However, participants in our study were merely instructed to specify their location rating for each word and did not perform any other language-related task (e.g., semantic judgement between word pairs), which would naturally trigger linguistic constructions (cf., Zwaan & Yaxley, 2003).

Thus, we conclude that location norms in our data set mostly depend on spatio-perceptual and bodily experiences associated with verbal representations. Along with that, spatial arrangements corresponding to certain abstract concepts still appear to be arbitrary. Conceptual metaphor theory claims that a pleasant person is "sweet" as sweetness is associated with pleasure at the sensory level (Lakoff & Johnson, 1980a; Meyer & Robinson, 2004). However, why goodness is associated with the upward position in the first place without any explicit metaphoric transfer remains as a critical question. Given that cultural differences (de la Fuente, Casasanto, Roman, & Santiago, 2015) or individual value systems (such as religiosity) (Chasteen et al., 2010) do not play a role in spatial mappings, possible answers might tap into the overarching questions of arbitrariness in language (Gasser, 2004; Pena, 2002; Seidenberg, 2002).

The main limitation of the study is the number of participants rated the words. A lab-based study enabled us to collect high quality data under controlled conditions yet with fewer participants. Although analyses indicated good external validity for vertical positions, results should be interpreted by taking the sample size into consideration. Further studies can adapt the coordinate paradigm for an online rating task in order to collect more participants and thus, bigger data. Variables such as linguistic, educational and sociocultural background, gender, body-specificities can be factored in with larger and more representative samples.

It is also important note that task instructions and the examples of spatial associations provided to the participants (see 5.8 Appendix) may have shaped participants' behaviour and thus, their ratings. Spatial associations in the current study should be interpreted considering this limitation. Although our findings are consistent with the previous studies exploring word-location associations where participants did not see explicit instructions before the test (e.g., Estes et al., 2015; Verges & Duffy, 2009), future norming studies could test spatial associations without any examples in the instructions.

We are of the opinion that the norms we provided in this study can be used to control and manipulate verbal material in studies exploring the relation between space and language. Main advantages of this data set are its size and precision and also, the availability of both horizontal and vertical norms for each word. In addition to the norms, the novel methodology we introduced paves a new way for norming spatial or spatially neutral words. Future studies can use the same paradigm to test different words, word types or words in different languages. The paradigm itself can also be improved. For example, the coordinate system can be integrated into Internet-based solutions such as Mechanical Turk by allowing to collect ratings from more people. Further, a three-dimensional coordinate system, measuring depth or the proximity of words to subjects in addition to horizontal and vertical positions, might shed further light on conceptual and spatial grounding of language.

5.7 Conclusion

The domain of language is critical for understanding grounded-embodied cognition and for any well-defined model that aims to explain sensorimotor operations. In this respect, spatial simulations evoked by language provide strong evidence for the sensorimotor origins of language within the framework of grounded-embodied cognition. Our study extends the literature of grounded-embodied cognition by providing further evidence for the link between language and spatial cognition. Word - space mappings does not seem to be arbitrary. Our results suggest that lexico-spatial associations are predominantly consistent, based on bodily/spatio-perceptual experiences and metaphorical groundings. Spatial simulations via language should have systematic implications for semantic processing even if space is irrelevant to the task considering the consistency of word - space mappings and that language-based simulations are triggered automatically.

5.8 Appendix

5.8.1 Instructions

Screen 1

Welcome to the spatial rating experiment. This experiment will take about 20 minutes. Please read the following instructions very carefully. Ask the experimenter if you have any questions at any point. Left click to proceed.

Screen 2

Words in language suggest the relative positions of the objects/concepts they stand for. For example, the word GRAVE suggests a relatively low location. Because a grave in the real world is located under the ground. Your task is to rate the locations of a number of words. This is done by left-clicking any point on a rating screen. Left click to see the examples of word and rating screens.

Screen 3

This is an example of the word screen. Please read the word carefully. Think about its meaning and try to visualise it. Then, press any key on the keyboard to see the rating screen. PLEASE TRY IT NOW!

Screen 4

(An example rating screen)

Screen 5

Everyone might have different word location concepts. That’s perfectly fine. There are no right or wrong answers in this experiment. We are simply asking YOUR opinion. For example, for some people, the word MAXIMUM might be located on the RIGHT end of the screen on a horizontal scale. For others, it might be located on the TOP of the screen on a vertical scale. Please use YOUR OWN intuition. Left click to proceed.

Screen 6

You will see some words, which might be located both on horizontal and vertical scales. For example, DAWN might be located BOTTOM LEFT if you think about the location of the Sun when it’s rising. Please think in all directions. Left click to proceed.

Screen 7

You will also see some words such as BALANCE, which might not make you think of any location at all. You can always click the central area to give these words somewhat a neutral location rating. This is also informative for us. Left click to proceed.

Screen 8

Rating screen has a very high resolution. It can distinguish tiny differences between click locations. You can click any point on the rating screen within the box. Please differentiate the words by considering this. For example, BUG might be located somewhat BELOW the previously seen word, COW. Left click to proceed.

Screen 9

You will see the following sentence and a box at the bottom right of the rating screen saying: I do not know this word well enough to give a rating. Left click the box if you don’t know the meaning of a word well enough to rate it. Please left click the box on the bottom right to proceed. PLEASE TRY IT NOW!

Screen 10

Now you will rate 11 words as a practice. If you have any questions, stop the practice and let the experimenter know. Left click to start the practice.

Screen 11

(Practice)

Screen 12

This is the end of the practice. Please let the experimenter know if you have any questions. The experiment will start when you left click and your ratings will be recorded. Please take your time, visualise the words and think carefully before rating. Thank you very much!

5.8.2 Words at the extremes of coordinate system

Table 5.6 Horizontal and vertical coordinates, controversiality and spatiality of the words located at the extreme positions of the coordinate systems

Location	Word	Horizontal	Vertical	Mean	Contraversiality	Spatiality
Top left	Northwest	-409.67	350.17	379.92	0.23	1.00
	Corner	-307.33	314.00	310.65	0.79	0.82
	Beginning	-366.67	133.50	250.09	0.41	0.66
	Tomahawk	-162.40	288.80	225.60	0.29	0.59
	Painter	-144.33	285.83	215.08	0.38	0.56
Top right	Justice	344.50	353.67	349.08	0.37	0.92
	Determination	309.50	382.83	346.17	0.35	0.91
	Excitement	246.83	399.50	323.17	0.27	0.85
	Climax	217.17	425.83	321.50	0.25	0.85
	Cure	320.00	288.17	304.08	0.39	0.80
Bottom left	Murder	-295.50	-427.50	361.50	0.32	0.95
	Suppression	-259.60	-444.20	351.90	0.28	0.93
	Grief	-326.33	-367.83	347.08	0.38	0.91
	Decadence	-276.00	-407.75	341.88	0.25	0.90
	Abasement	-218.00	-458.50	338.25	0.32	0.89
Bottom right	Arbiter	421.50	-244.50	333.00	1.00	0.88
	Nightfall	346.33	-167.67	257.00	0.47	0.68
	Whalebone	266.00	-98.00	182.00	0.28	0.48
	Rosin	203.50	-146.00	174.75	0.52	0.46
	Feudalism	186.00	-106.33	146.17	0.68	0.38
Right	Right	453.83	2.17	228.00	0.02	0.60
	Arbiter	421.50	-244.50	88.50	1.00	0.88
	Tomorrow	377.63	102.60	240.12	0.30	0.63
	East	360.83	11.17	186.00	0.23	0.49
	Nightfall	346.33	-167.67	89.33	0.47	0.68

Up	Magnitude	54.60	452.80	253.70	0.11	0.67
	Peak	0.00	451.00	225.50	0.04	0.59
	North	8.17	441.17	224.67	0.04	0.59
	Tower	26.17	432.83	229.50	0.10	0.60
	Rocket	53.67	431.83	242.75	0.29	0.64
Left	Left	-471.67	2.00	-234.84	0.00	0.62
	West	-456.83	1.33	-227.75	0.02	0.60
	Past	-422.50	-111.83	-267.17	0.21	0.70
	Northwest	-409.67	350.17	-29.75	0.23	1.00
	Yesterday	-390.50	-65.33	-227.92	0.26	0.60
Down	Abasement	-218.00	-458.50	-338.25	0.32	0.89
	Down	0.00	-455.00	-227.50	0.02	0.60
	Hell	-66.40	-454.00	-260.20	0.27	0.68
	Death	-87.33	-452.83	-270.08	0.22	0.71
	Suppression	-259.60	-444.20	-351.90	0.28	0.93

5.8.3 Excluded words

Table 5.7 Word that were rated as unknown by at least four participants and thus, excluded from the analysis

Rated as unknown by four participants	Rated as unknown by five participants
Abbess	Allegory
Banality	Belfry
Charlatan	Bivouac
Connoisseur	Blunderbuss
Debacle	Crag
Domicile	Deluge
Edifice	Encephalon
Fiord	Epistle
Firmament	Exactitude
Inanity	Gadfly
Increment	Gaiety
Indignation	Inclemency
Insolence	Kerchief
Kine	Loquacity
Osculation	Sonata
Prairie	Surtax
Racketeer	Titbit
Rheumatism	Thicket
Serf	Vestibule
Steerage	
Supplication	
Temerity	
Timepiece	
Warbler	

Chapter 6

Experiment 4

Simulating Space with Language: Effects of Congruency between Physical Space and Semantic Space on Memory for Language

6.1 Motivation and Aims

Two types of spatial simulations were examined in the previous chapters. Word locations were simulated in Experiment 1 and 2 when participants were asked to retrieve the words from memory in their absence. In Experiment 3, spatial norms indicated the potential locations of language-based spatial simulations based on either perceptual (e.g., “bird” - upward location) or metaphorical (“murder” - downward location) mappings. This chapter aims to explore how perceptual simulation of locations and language-based simulation of space interact in a single task and how these interactions affect retrieval performance in memory for language.

6.2 Abstract

Compatibility between the position of a word on a screen (i.e., physical space) with the position it implies as in “bird” = upward position (i.e., semantic space) can facilitate word processing (Zwaan & Yaxley, 2003). Further, words with such spatial associations can guide eye movements to the suggested locations and facilitate or interfere with the subsequent visual processing in the compatible location (e.g., detecting a visual target in upward position after reading the word “bird”) (Gozli et al., 2012). However, it is unclear how these two spatial dimensions (i.e., physical and semantic) interact in a single task. In the present study, we investigated the effect of spatial position during encoding and cueing/looking position during retrieval on memory performance for words with implicit spatial meaning or metaphorical spatial associations. Participants were asked to encode words presented in congruent (e.g., “bird” in an upward location) and incongruent locations (e.g., “bird” in a downward location). At retrieval, an auditorily presented word probed participants’ recognition memory as they were cued to look at either the original location of the probe word or a diagonal location. In

contrast to spatial congruency advantage in conceptual tasks, words presented in incongruent locations were retrieved faster than those in congruent locations. Further, being cued to congruent locations during retrieval improved memory performance for words encoded in incongruent locations but damaged the performance for highly imageable words encoded in congruent locations. Results provide strong evidence for the persistence of spatial simulation via language and indicate the role of physical space and mental imagery in spatial simulations via language.

Highlights

- A mismatch between a word's location on the screen and the location it implies (e.g., seeing "bird" at the bottom of the screen) led to better memory performance.
- Language-based simulations guided participants' eye movements to the implied locations. Seeing a visual cue in the compatible location interfered with memory and delayed the retrieval.

6.3 Introduction

"Ms. Cooper reached for her cup to have a sip of freshly brewed coffee. It smelled bitter but somehow fruity, which reminded her of the orange trees from her childhood." When readers of fiction read about the perceptual experiences of a character in a story such as smelling a cup of coffee, they tend to take the perspective of the character and have the same experiences as if they were real (e.g., Avraamides, 2003). Processing linguistic information (e.g., words describing the scent of coffee) activates perceptual information (e.g., scent of coffee). This phenomenon is known as *mental simulation* within the scope of grounded-embodied cognition (Zwaan, 1999).

Spatial perception, the perception of oneself and other objects in space, their locations and the spatial relations between them, can be simulated with language as well (see Speed & Vigliocco, 2015 for a review). For example, reading the word "bird" elicits a mental image which can give rise to a perceptual experience of "upward location" based on the typical location of a bird in the physical world (Bergen, Lindsay, Matlock, & Narayanan, 2007; Richardson, Spivey, Barsalou, & McRae, 2003). Further, abstract words with valence trigger simulations of metaphorical locations in space they are associated with (e.g., "justice" - upward location) (Chasteen et al., 2010; Leitan et al., 2015; Meyer & Robinson, 2004). Many concrete and abstract words in language, which we call *spatial words*, have such spatial associations (see Chapter 5) and thus, evoke spatial simulations.

The architecture of language-based spatial simulation is not well-defined. In particular, the literature is mixed with regard to the effect of physical space on language-based spatial simulations. First, there is evidence showing that spatial congruency between physical space and semantic space improves performance in conceptual tasks in which the response is made when the words are on the screen (e.g., Zwaan & Yaxley, 2003). Second, there is evidence that objects are identified more accurately when they are presented in congruent locations as to the locations implied by previously seen words (i.e., match advantage) (Chasteen et al., 2010; Meyer & Robinson, 2004; Zanolie et al., 2012) or more accurately when they are presented in incongruent locations (i.e., mismatch advantage) (Bergen, 2005; Bergen et al., 2007; Estes et al., 2008; D. C. Richardson et al., 2003).

The current study aims to shed light on the effect of physical location on spatial simulations evoked by language in memory. In particular, we tested the effect of congruency between physical space and semantic space at encoding and looking position at retrieval (congruent vs. incongruent as to the original location of the spatial word) on memory performance.

6.3.1 Physical and semantic space at encoding

Encoding in verbal memory involves the perception of words and generation of mental representations (Paivio, 1990). There is evidence that these representations do not only involve semantic information, but also information with regard to the location of words (see Spivey, Richardson, & Fitneva, 2004 for a review). Pylyshyn (1989) was among the first to propose a model (i.e., *spatial indexing*) in which locations are encoded along with the words themselves. In this model, a pre-attentive or a pre-cognitive mechanism in the visual system indexes the location of a visual stimulus by separating locations from other visual features (e.g., colour, shape etc.) even before recognising visual patterns.

Abundant evidence indicates that indexed locations interact with locations denoted by words upon presentation. For example, people react faster and more accurately in naming tasks if a word that conveys specific and explicit spatial information such as “left” appears in a congruent location on the display (i.e., left side of the screen) (White, 1969). On the other hand, inconsistency between spatial representations results in slower and less accurate responses (see Lu & Proctor, 1995 for a review). Known as *spatial Stroop*, this effect occurs even though the information about the word location is not necessary to complete the task successfully.

Reading words, which do not denote a location explicitly such as “left” but imply implicit locations such as “bird” at certain locations, has also consequences on how well these words are processed. Zwaan and Yaxley (2003) evidenced facilitation with spatial congruency in the case of words with imagined locations such as “basement” rather than the locations in the subject’s immediate environment. Participants were tasked to identify whether “root” and “branch” are semantically related. Reactions were faster and more accurate if “root” was located under “branch” compared to the incongruent arrangement (i.e., “branch” under “root”). In a similar paradigm, participants in Šetić and Domijan (2007) were tasked to verify whether a given word refers to a flying or a non-flying animal. Words for flying animals were verified faster when they were presented at the top of the screen and words for non-flying animals were verified faster when they were presented at the bottom (Experiment 1). The same effect was observed when participants verified living or non-living objects that are associated with upward or downward positions, with a categorisation task that was not spatial (Experiment 2).

Words that make direct reference to spatial locations (e.g., “above”) and words that occur in imagined but typical locations (e.g., “bird”) are not the only words associated with locations in space. Emotionally charged words with positive or negative valence such as “hero” or “liar” activate spatial simulations as well based on metaphorical relationship between valence and vertical space (i.e., “good is up”, “bad is down”) (Lakoff & Johnson, 1980b). Drawing on the associations between emotional words and space, Meyer and Robinson (2004) demonstrated that positive words are evaluated faster (i.e., “Rate ‘hero’ on a scale from 1 (negative) to 5 (positive)) if they are presented at the top of the screen and negative words (e.g., “liar”) are evaluated faster if they are presented at the bottom. In a similar fashion, Giessner and Schubert (2007) showed that people evaluate leaders as more powerful if their picture are positioned higher in a vertical arrangement (see also Schubert, 2005). In the domain of time - space mapping, it was demonstrated that linguistic units that imply time (e.g., “yesterday”, “we will drive”, “next” etc.) were judged faster when their position (i.e., left or right of the screen) matched with the location of response button (i.e., left or right hand-side button on the keyboard). For example, participants were faster to evaluate whether the sentence “we will drive” refers to past or future if the sentence was presented on the right side of the screen and response button was on the right hand-side (Santiago et al., 2007).

These studies reveal two mechanisms regarding space and language: First, word locations cannot be ignored even when they are irrelevant to the task (see Hock & Egeth, 1970; Seymour, 1977). In other words, individuals encode word locations automatically along with the words (see Chapter 3). Second, reading words with implicit spatial meaning simulates these locations automatically. When these two automatic mechanisms take place at the same time, a compatibility between the physical and sematic location of a word facilitates the response in tasks that require conceptual processing; whereas a conflict between these locations delays the processing.

Along with that, how simultaneous activation of physical and semantic representations of space affects memory performance is not clear. As a first question, we tested whether spatially congruent encoding (e.g., encoding "bird" at the top of the screen) results in faster retrieval in comparison to spatially incongruent encoding (e.g., encoding "bird" at the bottom of the screen).

The intuitive prediction in line with the studies discussed above is that spatial congruency at encoding leads to better memory. Under this prediction, participants are expected to process spatially congruent words faster and in contrast, spatially incongruent words slower. There is evidence to assume that encoding difficulty results in impaired memory as frequently seen in schizophrenia (M. J. Smith, Gorman, Malaspina, Sharif, & Amador, 2000) and older populations (Grady et al., 1995). There is also evidence showing that faster encoding usually results in faster and better retrieval (Kent & Lamberts, 2006; Lamberts, Brockdorff, & Heit, 2002). However, whether the ease of processing due to spatial congruency is robust enough to determine retrieval performance is not obvious.

The second prediction is that spatial incongruency rather than congruency will result in superior memory. This prediction might sound paradoxical; however, memory retrieval is fundamentally different than semantic judgement tasks reviewed above. Distinctiveness, in particular, has a substantial impact on verbal memory (see Hunt & Worthen, 2006 for a comprehensive review). Distinctiveness in the present study can occur and impact retrieval performance in two related ways:

Hirshman (1988) demonstrated that weakly related word pairs (e.g., quick - short) are better remembered than strongly related pairs (e.g., long - short). Weakly related word pairs are more likely to represent unexpected or novel semantic combinations and allow more elaborate encoding than expected pairs. In turn, elaborate encodings give rise to more accurate retrieval. In the current study, spatial words that are presented in incongruent locations can be assumed to violate participants' expectations due to spatial simulations. Consequently, spatially incongruent words might enjoy a similar memory advantage over spatially congruent words due to reconstruction of more detailed encodings.

Craik and Lockhart (1972) suggested that the memory of an item is a function of the depth to which the item was initially processed. In keeping with this argument, Craik and Tulving (1975) reported higher retrieval accuracy for words processed at deeper levels which were induced with semantic questions at encoding (e.g., “Does ‘apple’ fit into the ‘fruits’ category?”) as to words encoded with questions related to phonological (intermediate level) or orthographic (shallow level) properties of the word. We expect that participants encode all words at semantic level regardless of spatial congruency (see Chapter 4). That said, spatially incongruent words could be encoded at a relatively deeper level due to possibly additional resources (e.g., attention, cognitive effort) required to process an unexpected situation. In this respect, unexpectedness can lead to deeper and more elaborate encodings, which, in turn, could lead to a memory advantage for words encoded in incongruent locations.

In sum, evidence reported above demonstrates that words with spatial associations can give rise to simulations of concerned locations. Simulated locations interact with where the words are perceived and in return, either facilitate or delay the processing of spatial words. Previous studies suggest a congruency advantage in memory performance. However, incongruency between physical and semantic space at encoding could lead to better memory performance because unexpected items are typically remembered more accurately.

6.3.2 Physical and semantic space at retrieval

Pylyshyn's (1989) spatial indexing model posits that spatial indices persist across time, which makes it possible to locate and examine the encoded information if needed later on. Availability of spatial indices even when the information is no longer available manifests itself clearly during a memory retrieval behaviour, known as *looking at nothing* or *congruent area effect*. In looking at nothing, information to-be-retrieved is registered with spatial indices at encoding. When participants are asked to retrieve the information from memory in its absence, spatial indices attached to the information trigger eye movements to the location where the information appeared previously at encoding (e.g., Richardson & Spivey, 2000).

Semantic space persists in time as well. Simulated space (either through object words that denote certain locations or through abstract words that are grounded in certain locations) has consequences on the processing of subsequent visual information. This effect, known as *spatial cueing,* results in both facilitation and inhibition. For example, in Estes, Verges and Barsalou (2008), participants saw words that denote objects occurring in typical locations (e.g., "hat") in the centre of the screen. Later, participants were asked to detect unrelated visual targets (X or O) appearing either in the typical location of the object denoted by the preceding word (e.g., X on the upward location following "hat") or in an incongruent location (e.g., X on the downward location following "hat"). Visual targets appearing in the incongruent locations as to simulated locations (e.g., X on the downward location following "hat") were detected faster than visual targets appearing in the congruent locations (e.g., X on the upward location following "hat"). Participants had to supress the simulated location of the object (e.g., hat) to process the perceptual target occurring in the same location (e.g., X). This suppressing led to inhibition when identifying symbols in overlapping locations.

On the other hand, Chasteen, Burdzy and Pratt (2010) found that participants detected visual targets in congruent locations faster when their attention was guided by religious concepts associated with spatial locations (i.e., “God is up and right”, “Devil is down and left”). Response times were faster when targets appeared at the top and the right side of the screen following the concepts of God and when targets appeared at the bottom and the left side of the screen following the concepts of Devil (see also Amorim & Pinheiro, 2018; Zanolie et al., 2012). There is also evidence showing similar congruency advantage after reading words with positive or negative valence (e.g., processing a downward target following the word “tyrant”) (Zanolie et al., 2012) and retrospective or prospective time words (e.g., processing a leftward target following the word “tomorrow”) (Weger & Pratt, 2008).

Additionally, event-related brain potentials in Zanolie et al. (2012) revealed that words of power cause a shift in spatial attention towards specific positions in space in a time window from 160 to 200 ms. In a replication attempt of Chasteen et al. (2010), Gozli, Chasteen and Pratt (2012) reported early interference (short SOA between word and visual target = 200 - 400 ms) and late facilitation (long SOA between word and visual target = 800 - 1200 ms) and concluded that spatial cueing by words leads to two temporal stages of the same process: early interference and late facilitation.

The abovementioned evidence suggests that word locations are simulated in the absence of words and spatial simulations triggered by words persist in time and extend into the subsequent processing. How these mechanisms interact each other and in return, affect memory for language? An answer to this question could be informative as to the persistence of language-based spatial simulations and the effect of physical space on spatial simulations in the absence of words. To this end, participants in the current study are instructed to look at visual cues at the encoding stage appearing either in the congruent location as to the original location of the probe word or in a diagonal location.

Under the light of the studies reviewed above, two different effects can be observed in the current study. If spatial simulations evoked by words become decisive in memory performance, participants could treat the visual cue to orient their eye movements as additional visual information. In such a case, processing visual information in the same location as to the probe word could interfere with and delay the memory process (Estes et al., 2008) or alternatively, improve it (Chasteen et al., 2010; Gozli et al., 2012). However, if physical locations become decisive in memory performance, then visual cues could emphasize the previous locations of the words. In this case, a compatible visual cue would improve memory performance as in studies showing better memory performance with looks to previous locations of words (Johansson et al., 2012; Johansson & Johansson, 2014; Scholz et al., 2018, 2016).

6.4 Method

6.4.1 Participants

The experiment was carried out with forty-eight students at the University of Birmingham (five males; $M_{\text{age}} = 20.15$, $SD = 3.27$). 79% of the participants were psychology students. All participants were native speakers of British English as determined with the Language History Questionnaire (version 2.0; Li, Zhang, Tsai, & Puls, 2013). Participants reported normal or corrected-to-normal vision, no speech or hearing difficulties and no history of any neurological disorder. They received either £8 or course credit for participation. All participants were fully informed about the details of the experimental procedure and gave written consent. Post-experiment debriefing revealed that all participants were naïve to the purpose of the experiment.

6.4.2 Material

There were 192 trials involving 864 unique nouns in total. Trials were evenly divided into two groups (n = 96) as experimental (positive probe) trials and fillers. Probe words in the experimental trials were among the four study words in the encoding phase, whereas a different, not seen, word was probed in fillers.

Words in the experimental trials and fillers were arranged into sets with four and five words, respectively. Words within sets were matched on imageability, concreteness and length in letters (all *SD*s < 2.00). Words were further controlled so that no word started with the same letter or was semantically related with any other word in the trial set. Monosyllabic, disyllabic and trisyllabic words were evenly distributed [e.g., (3, 3, 3, 3), (1, 2, 1, 2) or (3, 2, 3, 2, 3) etc.]. Welch's t-tests revealed no significant difference between the probe and distractor words in imageability, concreteness, length in letters or number of syllables (all *p*s > .05). Thus, any word among the four words in each trial set was as likely to be remembered as any other word.

Probe words in experimental trials (n = 384) were selected from normed and labelled words according to their horizontal and vertical positions in Experiment 3. They were also a subset of the extension of Paivio's (1986) norms (J. M. Clark & Paivio, 2004). Equal number of words (n = 24) in the stimulus set were associated with top, top right, bottom and bottom left locations in space (see Chapter 5).

Horizontal positions were not considered as a variable due to the dominance of vertical space over horizontal space (Franklin & Tversky, 1990; Marmolejo-Ramos et al., 2013). Thus, words were grouped as upward (top and top right) and downward words (bottom and bottom left). All participants saw an equal number of upward and downward words (n = 48). Along with that, words were placed into quadrants according to both horizontal and vertical positions.

In the spatially congruent condition, words associated with upward and rightward position (e.g., "friend") were placed in the top right quadrant and words associated with upward position (e.g., "heaven") were placed in the top left quadrant. Similarly, words associated with downward and leftward position (e.g., "jail") were placed in the bottom left quadrant and words associated with downward position (e.g., "mule") were presented in the bottom right quadrant. In the spatially incongruent condition, words were encoded in the diagonal position as to their default position in congruent encoding. For example, participants saw a word associated with an upward position (e.g., "flag") in the top left quadrant when it was spatially congruent and, in the bottom right quadrant when it was spatially incongruent.

Finally, we formed 192 unique mathematical equation [e.g., (2*3) - (2+3) = 1] to present as memory interference between encoding and retrieval phases (see Conway & Engle, 1996 for a similar design). Half of the equations were correct. Incorrect equations were further divided into two equal groups: The results were either plus or minus one from the correct result.

Descriptive and spatial statistics of words used as probe are presented in Table 6.1 and Table 6.2 respectively.

Table 6.1 Descriptive statistics of probe words (n = 96)

Variable	Mean	SD	Minimum	Maximum
Imageability	4.93	1.34	2.20	6.77
Concreteness	4.43	1.96	1.18	7
Length in letters	6.76	1.80	4	11
Number of syllables	2.23	0.96	1	5
Horizontal position	-1.9	175.47	-326.33	344.5
Vertical position	2.51	321.65	-452.83	452.8

Table 6.2 Spatial statistics of the probe words (n = 96)

	Upward words				Downward words			
	Top		Top right		Bottom		Bottom left	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Horizontal position	33.87	75.82	221.74	52.83	-39.53	80.77	-226.81	48.32
Vertical position	327.84	68.41	289.35	74.99	-316.03	51.54	-317.66	72.74
Contraversiality	0.28	0.1	0.36	0.1	0.36	0.11	0.34	0.1
Spatiality	0.52	0.1	0.65	0.12	0.51	0.09	0.72	0.13

6.4.3 Apparatus

Stimuli were presented on a TFT LCD 24-inch widescreen monitor operating at 144 Hz with a resolution of 1920 x 1080 pixels (508 mm x 285.75 mm). The experiment was programmed in and run on OpenSesame (version 3.1.7; Mathôt, Schreij, & Theeuwes, 2012). Auditory material was produced by a native female speaker of British English in a sound attenuated room and recorded using Audacity (Version 2.1.10; https://www.audacityteam.org). Participants responded (yes/no they had seen the word) by pressing one of two keys on a standard keyboard.

Data were analysed and visualised in R programming language and environment (R Core Team, 2017). Mixed-effects models were constructed with *lme4* package (Bates et al., 2015). Significance values of the coefficients in models were computed based on the t-distribution using the Satterthwaite approximation with *lmerTest* package (Kuznetsova et al., 2015).

6.4.4 Procedure

The task was a yes/no verbal recognition memory test. As spelled out in detail below, each trial was composed of five consecutive phases: (1) fixation, (2) encoding, (3) interference, (4) retrieval and (5) feedback (See Figure 6.1). The task was to decide whether an auditorily presented word had appeared before or not. As soon as the participants made yes/no judgement by hitting one of the response buttons, the trial ended, and a new encoding phase began.

(1) Fixation: A fixation cross appeared at the centre of the screen for 500 ms. **(2) Encoding**: Participants were presented four words (Courier, font size = 56) in capital letters on a 2 x 2 grid for 1800 ms. Encoding location was a between-subjects variable. A random half of the participants (n = 26) saw the probe word in a spatially congruent location and the other half in an incongruent location. **(3) Interference:** Participants were presented a mathematical equation and asked to identify whether the equation was correct or not within 10,000 ms (or they timed-out). **(4) Retrieval**: The probe word was auditorily presented as participants looked at the screen with the blank grid. Participants were tasked to make a yes/no judgement to determine whether they had seen the probe word among the four words shown in the encoding phase within 3,000 ms (or they timed-out). Participants were asked to respond as quickly and as accurately as possible. During the retrieval phase, a square appeared either in the same (congruent retrieval) or in the diagonal quadrant (incongruent retrieval) as the original location of the probe word in the encoding phase. Participants were instructed to look at the square as they gave their response. Retrieval manipulation was a within-subjects variable. All participants were asked to retrieve the probe word in congruent and incongruent retrieval conditions. **(5) Feedback:** Participants were shown their accuracy (“correct” or “incorrect”) and response time for 500 ms after each trial. Total accuracy and average response time were shown at the end of the experiment.

The order of trials and equations were fully randomised independently of each other. The experiment was divided into four equal blocks with 48 trials in each block and there was a short pause between the blocks. A typical session lasted approximately 45 minutes. Overall accuracy in interference equations and in the recognition memory test for words (including fillers) were

79% and 81% respectively, suggesting that participants attended to the task with high concentration.

Following the experiment, a computerized version of the Corsi block-tapping task (Corsi, 1972) operated on PEBL (Psychology Experiment Building Language, version 0.13, test battery version 0.7, http://pebl.org) (Mueller & Piper, 2014) was used to measure visuospatial short-term memory and a digit span test was used to measure short-term memory. Participants were also asked to report their level of sleepiness on Epworth Sleepiness Scale (Johns, 1991). Mental fitness was measured on a Likert scale from 0 (“I feel very tired”) to 4 (“I feel mentally very fit and attentive”). Physical fitness score was computed based on exercise frequency. Psychological mood was measured on a Likert scale from 0 (extremely unhappy) to 10 (extremely happy).

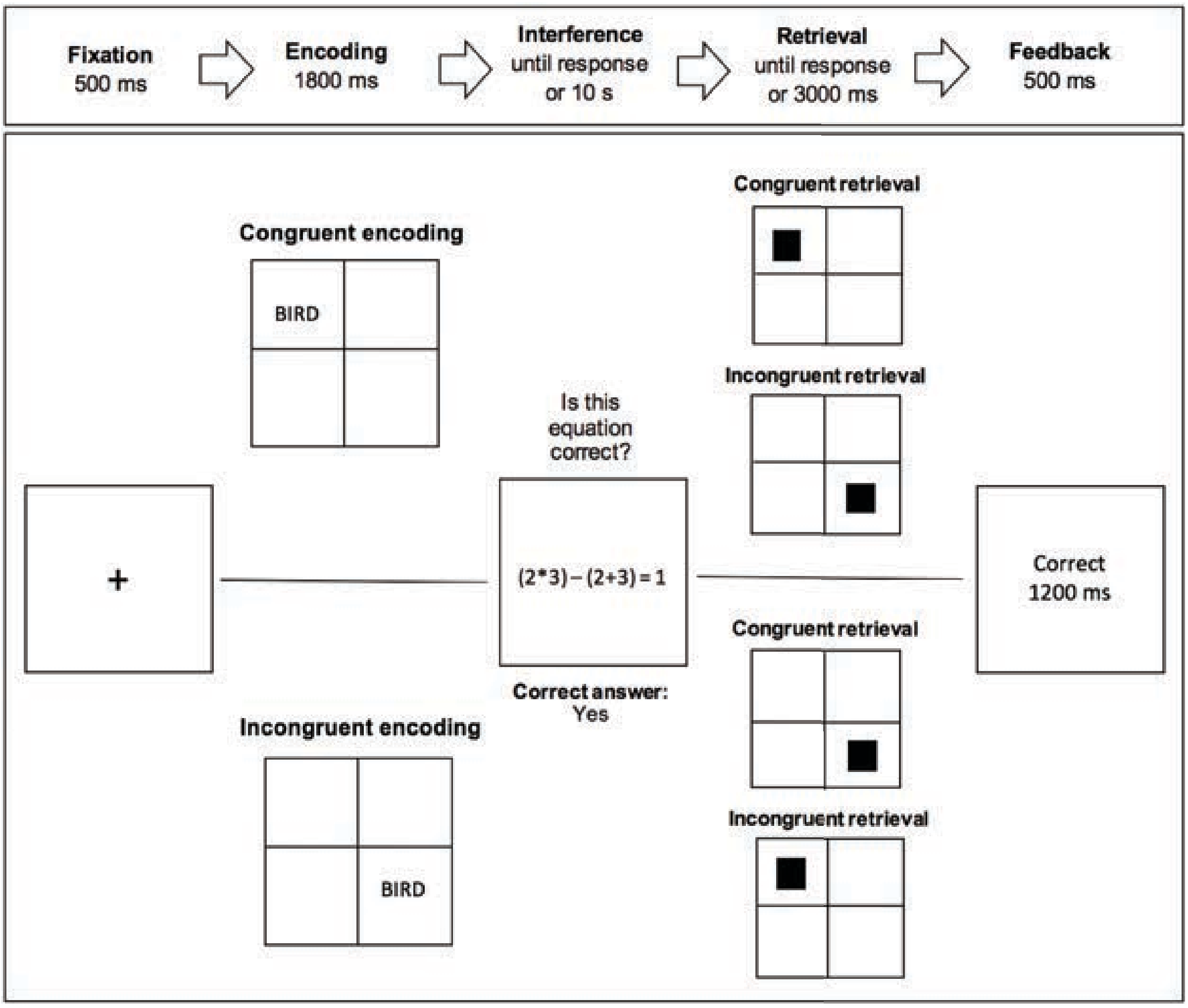

Figure 6.1 A schematic illustration of the temporal order of events in an example trial showing two different encoding and two different retrieval conditions. In this example, the relevant quadrant is the top left location, where the probe word (i.e., BIRD) appears.

6.5 Results

6.5.1 Data preparation

Homogeneity

Variability between the participants was considered as a random effect in the mixed-effects models. Along with that, participants who were tested on congruent and incongruent encoding conditions were compared on individual variability measures (i.e., Corsi Block-tapping test, digit span, Edinburg Handedness Inventory, Epworth Sleepiness Scale, self-reported mental fitness, self-reported physical fitness and self-reported mood) to verify the homogeneity between the groups. There was not a significant difference between participants tested on congruent and incongruent encoding conditions in any of the measures (see Table 6.3).

Table 6.3 Results of the individual variability measures among all participants and across encoding groups

Measures	All (48)	Congruent encoding (24)	Incongruent encoding (24)	*t*	*p*
Corsi block	58.94 (24.25)	53.75 (25.01)	64.12 (22.81)	1.50	.14
Digit span	7.06 (1.36)	9.62 (2.58)	8.67 (2.16)	1.39	.17
Handedness laterality	59.17 (49.17)	54.79 (49.35)	63.54 (49.64)	0.61	.54
Sleepiness	0.84 (0.55)	0.85 (0.51)	0.83 (0.6)	0.13	.90
Mental fitness	2.21 (0.92)	2.25 (0.94)	2.17 (0.92)	0.31	.76
Physical fitness	3.21 (1.52)	3.21 (1.61)	3.21 (1.44)	0	1
Mood	6.85 (1.22)	6.62 (1.13)	7.08 (1.28)	1.31	.20

Measures

4608 observations from 48 participants (excluding fillers) were included in the hit rate analyses and 3721 observations from 48 participants (excluding fillers and incorrect responses) were included in the hit latency analyses. One participant was replaced due to low accuracy in the recognition memory test (46%) below our threshold (60%). Total accuracy in the experimental trials after excluding one participant was 80.75% (*SD* = 0.39).

Hit rate and hit latency were used as measures of memory performance. Hit rate was the proportion of experimental trials to which the participants correctly responded yes. Hit latency was the time in milliseconds between the onset of auditory presentation of the probe word and correct keyboard response. Exceptionally slow hit latencies that are 2 SD away (+/- 678.5 ms) from the mean (1079.41 ms) were omitted (*n* = 179, 4.81%) from the hit latency analyses. Overall, 3542 observations from 48 participants were analysed for hit latencies. Densities of hit latencies before and after trimming are presented in Figure 6.2.

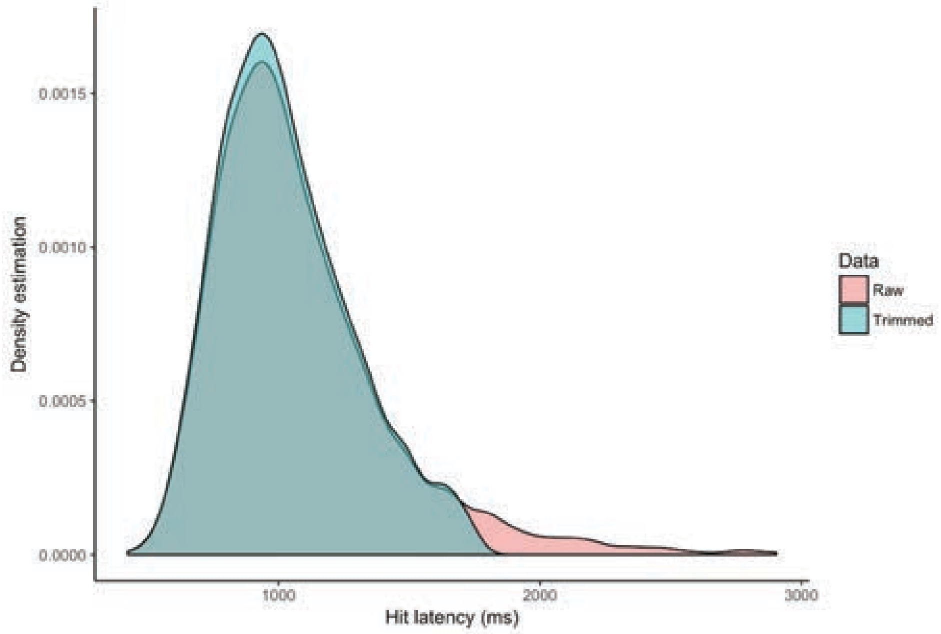

Figure 6.2 Density plot showing distributions of hit latencies before and after data trimming.

6.5.2 Mixed-effects modelling

Data were analysed using linear and binomial logit mixed-effects modelling. Visual inspections of residual plots did not reveal any obvious deviations from homoscedasticity or linearity. Linear models were fit for continuous target variables (hit latency). Binomial models were fit for categorical target variables (hit rate) and with *bobyqa* optimiser to prevent non-convergence. Participants and items were treated as random effects to explain by-participant and by-items variation (Baayen, Davidson, & Bates, 2008).

We started fitting models by building the random effects structure and followed a maximal approach. That is, random effects were included as both random intercepts and correlated random slopes (random variations) as long as they converged and were justified by the data (Barr, Levy, Scheepers, & Tily, 2013). Random intercepts and slopes were included even if they did not improve the model fit in order to control for possible dependence due to repeated measures or order effects. Random effects structure was simplified step by step as per the magnitude of the contribution of a random effect to the explanation of the variation in the data. That is, the random effect with the weakest contribution was dropped first and if necessary, the structure was further reduced accordingly.

Contribution of a fixed effect was investigated by comparing a full model containing the effect in question against a reduced model in which only that effect was removed, or a null model without any fixed effects. Compared models had the same random effects structure (Winter, 2013).

6.5.3 Factor analysis

Diagnostic tests indicated collinearity between word variables ($Mean_{VIF}$ = 2.86, $Range_{VIF}$ = 2.06 - 3.46; kappa = 2400.33). In order to address collinearity, we performed exploratory factor analysis using a principal component analysis extraction method with an orthogonal (varimax) rotation method. Kaiser's criterion and the scree test criterion showed the presence of three factors in our data. Three factors were interpreted as follows based on the loadings (see Table 6.4): **(1) Imagery**[5]: Imageability and concreteness, **(2) Length:** Length in letters and syllable length and **(3) Position:** Horizontal and vertical position. The three-factor solution explained 89.38% of the variance in the data. VIF of the factors were below one and thus, below our threshold of two. Regression scores calculated for three factors were employed as predictors in the subsequent linear mixed-effects multiple regression models.

Models were fit to investigate the effect of word factors on hit latency. Random slopes were added for length and imageability into items and length into participants. Likelihood tests showed that word length (length in letters and syllable length) was predictive of hit latency; $\chi^2(1) = 3.76$, $p = .05$ but not imageability; $\chi^2(1) = 1.49$, $p = .22$ or word position; $\chi^2(1) = .007$, $p = .93$. Longer words were retrieved slower; $\beta = 14.22$, $t = 2.17$, $p = .04$.

[5] We refer to imagery factor (i.e., imageability and concreteness) as imageability for the sake of simplicity.

Table 6.4 Varimax rotated factor-loadings and communalities of the predictors

	Factor			
Predictors	Imagery	Length	Position	h^2
Imageability	**0.93**	-0.21	-0.02	0.92
Concreteness	**0.95**	-0.15	-0.06	0.92
Length in letters	-0.19	**0.93**	0.11	0.91
Number of syllables	-0.17	**0.94**	0.05	0.91
Horizontal position	-0.02	0.09	**0.92**	0.85
Vertical position	-0.05	0.06	**0.92**	0.86
Factor statistics				
Eigenvalue	2.62	1.60	1.13	7.99
Variance (%)	30.55	30.28	28.54	89.93

Bolded numbers indicate the groupings. Eigenvalues and percentage of variance are after rotation. h^2 = communality.

6.5.4 Effect of spatial congruency at encoding

Hit rate

First, we examined the effect of spatial congruency between word locations and locations implied by the words on hit rate in the verbal recognition memory test. Fixed effect was encoding condition with two levels (spatially congruent encoding and spatially incongruent encoding). Random slopes were added for imagery, word length and word position factors both into participants and items. Likelihood tests showed that encoding condition did not improve the model fit for hit rate; $\chi^2(1) = 0.49$, $p = .48$. In other words, participants' accuracy was the same for words encoded both in canonical (mean hit rate = 80%) and unexpected locations (mean hit rate = 82%). None of the lexico-semantic factors modulated the effect of spatial congruency at encoding on hit rate (ps > .05) (see Figure 6.3).

Hit latency

Next, we examined the effect of spatial congruency at encoding on hit latency. Fixed effect was encoding condition with two levels (spatially congruent encoding and spatially incongruent encoding). Random slopes were added for imagery, word length and word position factors both into participants and items. Likelihood tests showed that encoding condition significantly improved the model fit; $\chi^2(1) = 6.96$, $p = .008$. Words encoded in congruent locations (e.g., “bird” in the top left quadrant) (mean hit latency = 1075.26 ms) were retrieved slower than words encoded in incongruent locations (e.g., “bird” in the bottom right quadrant) (mean hit latency = 985.63 ms); $\beta = 95.29$, $t = 2.80$, $p = .007$. None of the lexico-semantic factors modulated the effect of spatial congruency at encoding on hit latency (*p*s > .05) (see Figure 6.3).

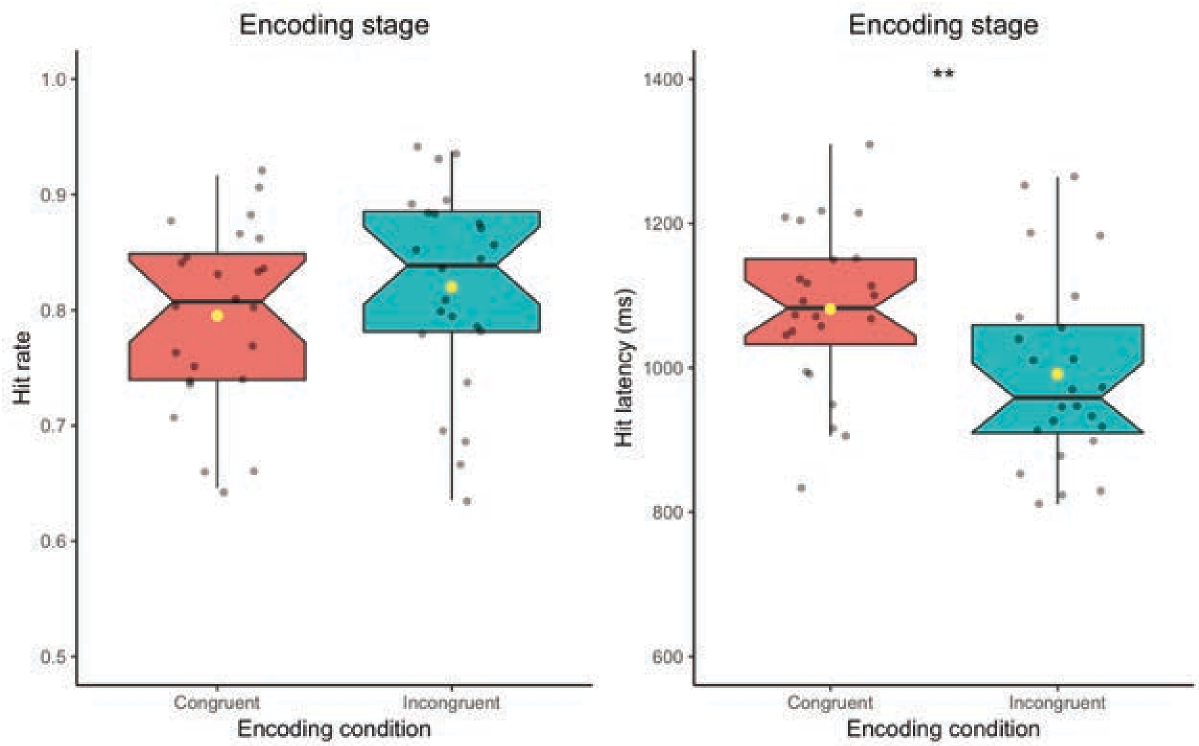

Figure 6.3 Hit rate and hit latencies across spatially congruent and spatially incongruent encoding conditions. Notched box plots show median (horizontal line), mean (yellow dot), 95% confidence interval of the median (notch), interquartile range (the box), the first and the third quartiles (lower and upper ends of the box) and ranges (vertical line). Grey dots represent data points. ** $p \leq .01$

6.5.5 Effect of spatial congruency at retrieval

We fit separate models within congruent encoding and incongruent encoding conditions to examine the effect of cueing/looking position at the retrieval stage. Fixed effect was retrieval condition (i.e., the location of the visual cue presented at retrieval) with two levels (spatially congruent retrieval and spatially incongruent retrieval).

Hit rate

Congruent encoding: Word length factor was added as random slopes into participants and items. Retrieval condition did not improve the model fit for hit rate within congruent encoding condition; $\chi^2(1) = 1.23$, $p = .27$. In other words, participants were equally accurate in word retrieval when they were cued to look at the previous locations of the probe word (mean hit rate = 79%) or at a diagonal location (mean hit rate = 80%). None of the lexico-semantic factors modulated the effect of spatial congruency at retrieval on hit rate within congruent encoding trials (*p*s > .05).

Incongruent encoding: Word length and word position factors were added as random slopes into participants and items. Retrieval condition improved the model fit for hit rate within incongruent encoding condition; $\chi^2(1) = 4.59$, $p = .03$. Participants were more accurate in word retrieval when they were cued to look at the previous locations of the probe word (mean hit rate = 84%) than to a diagonal location (mean hit rate = 80%); $\beta = 0.25$, $z = 2.11$, $p = .03$. Findings are visualised in Figure 6.4.

Hit latency

Congruent encoding: Imagery and word length factor were added as random slopes into participants. Word length factor was added into random slopes into items. Retrieval condition did not improve the model fit for hit latency within congruent encoding condition; $\chi^2(1) = 0.10$, $p = .75$. Participants were equally fast in word retrieval when they were cued to look at the previous locations of the probe word (mean hit latency = 1081.80 ms) or at a diagonal location (mean hit latency = 1082.01 ms).

There was a significant interaction effect between retrieval condition and imageability on hit latency; $\chi^2(1) = 5.58$, $p = .02$. Participants retrieved high imageable words slower when they read the word in the congruent locations and then, were cued to look at the previous location of the word; $\beta = 25.32$, $t = 2.36$, $p = .02$.

Incongruent encoding: Imagery, word length and word position factors were added as random slopes into participants. Imagery factor was added as random slopes into items. Retrieval condition did not improve the model fit for hit latency within incongruent encoding condition; $\chi^2(1) = 2.42$, $p = .12$. Participants were equally fast in word retrieval when they were cued to look at the previous locations of the probe word (mean hit latency = 999.04 ms) or at a diagonal location (mean hit latency = 983.06 ms). There was not an interaction effect between retrieval condition and any lexico-semantic variable on hit latency within incongruent encoding. Findings are visualised in Figure 6.4.

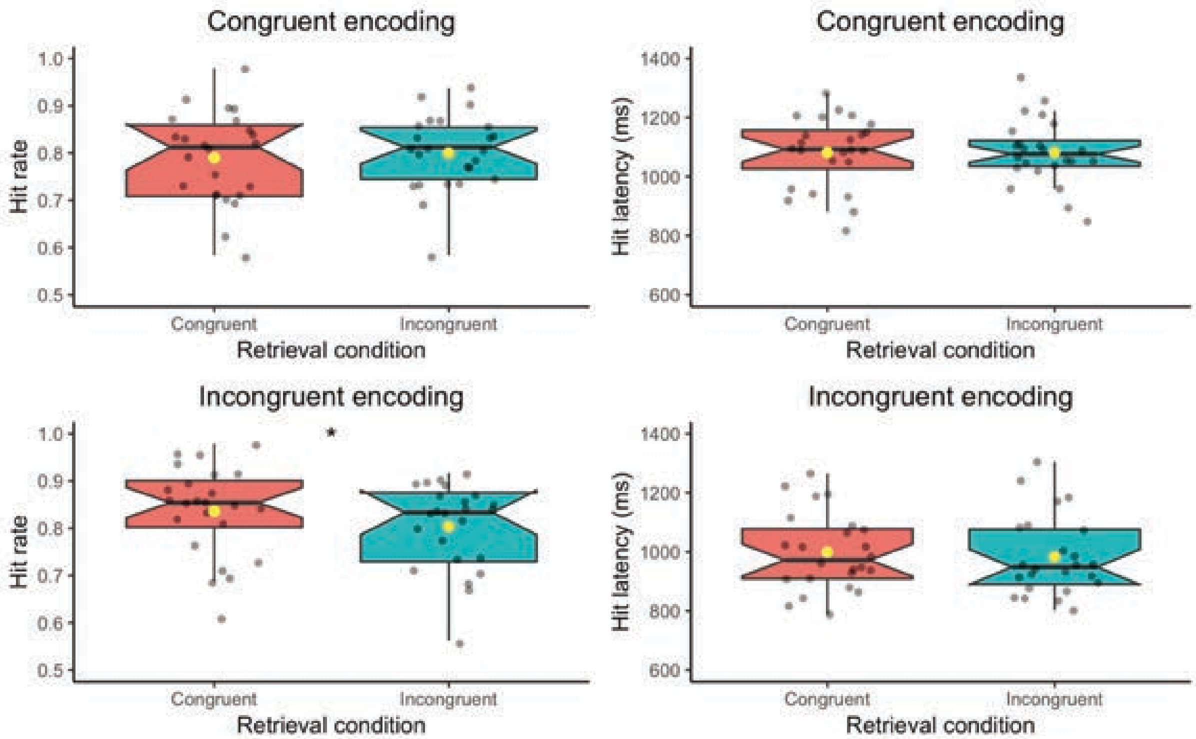

Figure 6.4 Hit rate and hit latencies across spatially congruent and spatially incongruent retrieval conditions within congruent and incongruent encoding conditions. Notched box plots show median (horizontal line), mean (yellow dot), 95% confidence interval of the median (notch), interquartile range (the box), the first and the third quartiles (lower and upper ends of the box) and ranges (vertical line). Grey dots represent data points. * $p \leq .05$

6.6 Discussion

We investigated compatibility effects of physical and semantic locations of words at encoding and retrieval stages on memory performance. Physical space was the actual position where the word was presented on the screen at encoding and semantic space was the location suggested by the word either perceptually or metaphorically.

6.6.1 Physical and semantic space at encoding

Results showed that words that were encoded in incongruent locations as to the locations they are associated with were retrieved faster. To illustrate, participants were faster when remembering the word "bird" if they saw it at the bottom of the screen as compared to the top of the screen. The effect was evident even when word properties (length in letters, syllable length, imageability, concreteness, horizontal and vertical positions) that have evidenced impacts on verbal memory under the current paradigm (see Chapter 4) were controlled as much as the models permit.

The results are in contrast with studies showing facilitation with spatial congruency in other linguistic tasks such as semantic judgement (Zwaan & Yaxley, 2003) and conceptual categorisation (Šetić & Domijan, 2007). However, verbal memory can be assumed to require different mental operations than other language/conceptual tasks as discussed in introduction. Consequently, memory-related effects seem to have played a more decisive role in retrieval performance in the present study. Thus, we refer to the research on memory, in particular, the concept of distinctiveness to explain our results. We assume that participants retrieved the words that appeared in unexpected locations faster as these words were encoded more distinctively. According to this account, spatial congruency might indeed have led to ease of conceptual processing. However, encoding in the current study was not a production but a perception task itself. That is, participants were not instructed to produce a respond at the processing stage unlike studies showing facilitation with spatial conformity (e.g., Meyer & Robinson, 2004; Šetić & Domijan, 2007; Zwaan & Yaxley, 2003).

We argue that expectation-violation effect (Hirshman, 1988) and level of processing (Craik & Lockhart, 1972) together could explain the spatial incongruency advantage in the present study. Experiment 3 indicated that there are strong and consistent associations between words and their simulated locations in space. Thus, we reason that participants in the current study had consistent spatial expectations as to the words they encoded. It is possible that the need to resolve the inconsistency between the word location and the location implied by the word resulted in deep encoding elaboration. In support of this account, there is a large body of evidence that such complex, elaborate encodings lead to better recognition memory by leaving richer traces (R. P. Fisher & Craik, 1980; Galli, 2014; Hunt & Worthen, 2006; Wagner et al., 1998). In sum, spatially incongruent words were paradoxically remembered faster in our study because such words violated spatial expectations and thus, encoded deeper than words that were presented in locations in line with their implied locations.

Findings from encoding condition can be regarded as expected within memory research. That said, they are noteworthy from the standpoint of simulation approach to language and thus, grounded-embodied cognition. In the current study, the embodied mind undertook two spatial operations at different dimensions (i.e., physical and semantic) at the same time: Participants registered locations of the words on the screen and the words activated spatial perceptions. Spatial information was neither relevant nor necessary for a successful retrieval in the present study. Even though, both external and internal spatial perception, that is, perception of word location on the screen and simulation of implied locations had effect on memory for language, which is seemingly a non-spatial and predominantly internal task. Importantly, we had observed the effect not only for concrete nouns that denote object occurring in typical locations in the real world but also for abstract nouns that are metaphorically associated with spatial locations. Such a finding suggests the pervasiveness of spatial perception in linguistic operations. Hence, results provide evidence for grounding of both concrete and abstract concepts in space and on a larger scale, for a simulation-based understanding of language (Barsalou, 1999; Bergen, 2007).

6.6.2 Physical and semantic space at retrieval

Visual cues presented in either the same or a diagonal location as to the original location of the probe word affected memory performance according to the initial location of the probe word at encoding:

(1) Participants who were cued to look at the previous location of a high imageable word encoded in a congruent location were slower than the participants who were cued to look at a diagonal location. To illustrate, responses were slower when participants were cued to look at the top left quadrant after seeing the word "bird" (but not "justice") in the top left quadrant.

(2) Participants who were cued to look at the previous location of a word encoded in an incongruent location were more accurate than the participants who were cued to look at a diagonal location. To illustrate, responses were more accurate when participants were cued to look at the bottom right quadrant after seeing the word "bird" (or "justice") in the bottom right quadrant.

Both findings indicate that words activated spatial simulations and guided participants' eye movements to the locations they suggest due to involuntary spatial cueing (e.g., looking upward after reading the word “bird”) (Dudschig et al., 2013). Later, seeing visual information in the compatible location interfered with the retrieval process (Bergen, 2005; Bergen et al., 2007; Estes et al., 2008; D. C. Richardson et al., 2003). In the first case, reading the word “bird” oriented eye movements to the top left quadrant and processing a visual cue in the same location delayed the retrieval process. In the second case, reading the word “bird” oriented eye movements to the top left quadrant again and processing a visual cue in a different location than the simulated location (i.e., bottom right quadrant) improved memory performance. In short, memory performance was better whenever the simulated location did not overlap with the cue location.

We assume that the interference effect was a consequence of the competition between the simulated location and the visual information when they occurred in the same quadrant. Although participants were not tasked to act on the visual information shown at retrieval and thus, they did not have to supress the simulated location to respond (cf., Estes, Verges, & Barsalou, 2008), they still had to allocate cognitive resources to overlapping physical and semantic spatial information. Hence, a competition for cognitive resources could have led to worse memory compared to the retrieval condition in which the simulated location and visual information did not overlap.

Memory advantage with spatial incongruency at retrieval is in line with studies showing mismatch advantage in object identification tasks (Bergen, 2005; Bergen et al., 2007; Estes et al., 2008; D. C. Richardson et al., 2003) but at odds with studies showing a match advantage due to spatial simulations (Amorim & Pinheiro, 2018; Chasteen et al., 2010; Zanolie et al., 2012). Notably, previous evidence suggests that simulation via concrete language leads to a match advantage; while, simulation via abstract language leads to a mismatch advantage (Liu & Bergen, 2016). We did not group words as concrete and abstract; however, it is possible that abstract language with high emotionality and thus, strong spatial associations (see Chapter 5) might have affected the results over concrete words in the present study.

Our results are also at odds with the evidence highlighting the functional role for eye movements in memory (Johansson & Johansson, 2014; Scholz et al., 2018, 2016). We argue that semantic space rather than physical space played a key role in memory retrieval when words suggest locations in space. As a result, participants treated visual cues as additional information to be processed rather than location tags that remind them the previous locations of the words. If this was the case, performance should have been better when participants looked at previous locations of words encoded in congruent locations.

The mediating role of imageability for words encoded in congruent location is consistent with this explanation. We assume that imageability exaggerated the spatial simulation and in turn, supported the interference at the simulated locations. Accordingly, "bird" (a highly imageable word) elicited a more vivid mental image and thus, gave rise to a stronger spatial experience as compared to a low imageable word such as "justice". Experiment 3 showed that low imageable words are also associated with certain locations in space according to their valence. We assume that reading these words activated their locations to a certain degree. However, this simulation was not robust enough for the interference effect when the words were encoded in congruent locations.

Why did words encoded in incongruent locations lead to interference effect without the mediating role of imageability but only high imageable words encoded in congruent locations resulted in interference? We speculate that words encoded in incongruent locations could have given rise to stronger spatial simulations due to the spatial unexpectedness discussed above. Put differently, an unexpected location could have emphasised the spatial meaning of a word and consequently, the interference due to spatial cueing was more pronounced for words encoded in incongruent locations.

One important issue to consider is that spatial manipulations in the current study affected accuracy or response time under different conditions. To be more precise, spatial congruency at encoding affected hit latency but not hit rate. Spatial congruency at retrieval affected hit latency for words encoded in congruent positions and hit rate for words encoded in incongruent positions. There is evidence that attentional cues affect accuracy and response time via different cognitive and neural processes (van Ede, de Lange, & Maris, 2012). That is, accuracy is explained by a single process (i.e., increase in preparatory excitability in the sensory cortex), whereas response time is explained by an additional process that is sensitive to the compatibility between cue and target (post-target comparison between expected and actual location of stimulus). This evidence suits well with the effect of spatial congruency at encoding on hit latency in that response time was more sensitive to the spatial manipulation. In another study, the effect of spatial simulation activated by emotional words was not reflected in accuracy data but only in reaction times (Amorim & Pinheiro, 2018). However, this account fails short to explain why spatial congruency affected hit latency for words encoded in congruent positions but hit rate for words encoded in incongruent positions. Further research disambiguating the effects of spatial simulations on hit rate and hit latency is necessary.

Results have a number of implications for the interactions between space, language and perceptual simulation. First, we provided further evidence that simulated locations can persist in time with consequences in the later stages of a task. The novelty of our contribution is the robustness of spatial simulation and persistency of semantic space in a longer time-window. Locations activated by spatial words in the current study affected subsequent processing even after participants solved a maths problem as an intermediate task. Thus, the duration between stimulus (i.e., spatial word) and response was 5351 ms on average in the current study. Such a result goes against the evidence suggesting that language-based simulation decays in the following 800 ms from the stimulus onset (Ostarek & Vigliocco, 2017) or that simulations have later effects within a time window of 800 - 1200 ms but congruency results in facilitation in the performance (Gozli et al., 2012). Language-based simulations had effects in a much longer time-window in our study and congruency resulted in interference. We argue that when participants retrieved the words, they retrieved the suggested locations associated with the words as well. We did not manipulate and investigate stimulus-onset asynchrony in the present study. However, our results suggest a potential direction for future studies to investigate the cueing effect of spatial words on the subsequent visual information processing at longer SOAs.

Second, we provided the preliminary evidence that spatial cueing could be stronger when spatial words are presented in different locations as to locations they imply. To our knowledge, spatial arrangement at the presentation has not been manipulated as a variable before in studies exploring subsequent effects of spatial simulations. Our results suggest that compatibility of semantic space with the physical space can enhance or dampen the effect of simulated locations. This also paves a new way for the future studies looking at the relationship between space and language.

Lastly, we showed that imageability as a word property and by extension, mental imagery is critical in perceptual simulation (see Kosslyn, Thompson, & Ganis, 2006). Mental imagery has been used in many studies as a representative operation to show the existence and reactivation of perceptual symbols in that perceptual symbols (e.g., eye movements) encoded at the viewing phase are simulated at the imagery stage (e.g., Laeng & Teodorescu, 2002; Verges & Duffy, 2009). Mental imagery as a function of word imageability appears to modulate the strength of perceptual experiences through linguistic processing (see also Experiment 2).

The current experiment has a number of limitations. First, we have no proof that participants actually looked at locations they were cued to as we did not track eye movements although experimental instructions and systematic effects of visual cues suggest that they did. That said, accompanying eye movement data in a future study could verify the effect of language-based simulations on visual cues. Second, two different effects were investigated in the present study: (1) the congruency between physical and semantic location of a word at encoding and (2) the congruency between physical/semantic location of a word and subsequent visual information. These two effects might have been intermingled and obscured each other when investigated in one experiment with a relatively complex design. A more elegant experiment that focuses on either the first or the second effect would further clarify the role of spatial simulation in memory for language. For example, in one experiment, participants can be instructed to remember spatially congruent or incongruent words as to the locations they imply without showing any cues at retrieval. In another experiment, participants can be instructed to encode words in the central (thus, neutral) position. At retrieval, they can be asked to retrieve the probe word while looking at a visual cue that is congruent or incongruent as to the physical or semantic word location at encoding.

6.7 Conclusion

Space in language was investigated at two dimensions in the present study: semantic and physical. A mismatch between semantic and physical space made words more distinct and thus, more memorable. Alternatively, the simulated semantic location interfered with subsequent processing in the congruent location and delayed retrieval. In that, our results allow for a more precise characterization of the nature of language-based simulations. Both cases underline the importance of space in language. We argue that spatiality as a word property could be taken into consideration in models of language processing and memory for language. Overall, results indicate that space is automatically simulated via language with highly persistent and pervasive consequences on mental operations such as verbal memory. Results can be taken as further support for a perceptual and extended understanding of language (e.g., Gibbs, 2005): Language is not an amodal, symbolic abstraction that is independent of other cognitive faculties in mind. Rather, it is a flexible system grounded in perceptual simulations and interacting with the physical environment.

Chapter 7

General Discussion

7.1 Summary of Main Findings and Conclusions

The current thesis explored and illustrated spatial simulations towards more precise simulation models within the framework of grounded-embodied theories of memory and language. Simulation of word locations in the absence of words during memory retrieval was addressed in Chapter 3 (Experiment 1) and Chapter 4 (Experiment 2). Simulation of spatial locations triggered directly and automatically by word meanings was focused in Chapter 5 (Experiment 3) and Chapter 6 (Experiment 4).

Experiment 1 results showed that participants encoded locations of words upon their presentation in an unintentional and effortless manner: First, word locations were encoded along with the words themselves although participants did not have to remember locations intentionally. Second, words locations were encoded without any objects (e.g., faces) to explicitly associate with as in previous studies (thus, effortless) (D. C. Richardson & Spivey, 2000; Scholz et al., 2018, 2011, 2016). Spatial indices tagging word locations induced simulation of the location when the information previously occurred there was re-accessed with memory retrieval. In return, simulation of word locations guided eye movements to the relevant locations (i.e., looking at nothing). In other words, participants relied on the simulation of word locations. In this manner, memory work was offloaded onto the body and the environment through eye movements and spatial simulations, respectively.

Experiment 1 results also showed that spatial indices can be modulated with additional visual information. A visual cue appearing in the same location as to the previous location of the probe word emphasised the locational information and strengthened its simulation. As a result, looks to the relevant, blank locations were more frequent following a spatially congruent visual cue compared to trials where no cue was shown. On the other hand, a visual cue appearing in a diagonal location to the previous location of the probe word formed a new spatial index. Competition between the spatial indices standing for the probe word and the incongruent visual cue interfered with "looking at nothing" behaviour but it did not delay memory retrieval.

The key findings of Experiment 1 are the correlations between visuospatial memory span measured with Corsi-block tapping task and looking behaviour: First, there was a positive correlation between visuospatial memory and percentage of fixations in the central interest area (across all cue conditions). Second, there was a negative correlation between visuospatial memory and percentage of fixations in the relevant, blank quadrant (in incongruent cue condition) and in the irrelevant, blank quadrant (in no cue condition). We conclude that participants with worse visuospatial memory relied more on the environment and thus, simulation of word locations than participants with better visuospatial memory.

Almost all patterns of the looking behaviour observed in Experiment 1 were replicated in Experiment 2: First, participants looked at the relevant, blank locations (the quadrant where the probe word was presented) more than the irrelevant, blank locations (the other three blank quadrants) when remembering words. Second, visual cues shown between encoding and retrieval stages affected looks in the same way as found in Experiment 1. Participants looked at "nothing" when no cue was presented or following a spatially congruent cue. However, spatially incongruent cues disrupted the looks to blank locations. Third, the link between visuospatial memory and looking behaviour was replicated. Lastly, looking at nothing did not improve memory performance in either of the studies.

What is the novel contribution of Experiment 2? The key finding is that participants looked more at "nothing" when remembering words that are more difficult to maintain and retrieve from memory compared to easier words. Mixed-effects models that were fit to investigate the contribution of lexico-semantic variables revealed that word imageability and concreteness modulated looks to blank locations. There were more looks in the relevant, blank locations during the retrieval of less imageable and abstract words as compared to more imageable and concrete words. Notably, word length modulated memory performance (hit rate and hit latency) but not looks to relevant, blank locations suggesting a special role for imageability and by extension, mental imagery. In support of this, participants with more control on their mental images as measured with Gordon Test of Visual Image Control (Gordon, 1949) relied on spatial simulations more frequently (i.e., looked more at nothing) than participants with worse mental imagery control when there was no visual cue between encoding and retrieval stages. The contribution of semantic variables (i.e., word imageability and concreteness) but not perceptual variables (i.e., word length) indicates that participants did not treat words as physical images. Rather, they accessed word meanings and a semantic component modulated looking behaviour. Such a behaviour highlights the tight links between spatial and conceptual representations (Boot & Pecher, 2011; Martarelli et al., 2017).

We attempted to specify the mechanism of location simulation in memory for language in Experiment 1 and 2. In particular, our evidence illuminated the conditions under which individuals tend to rely on location simulations and look at relevant, blank locations during memory retrieval. We conclude that there is a balanced trade-off between internal and external sources driven by cognitive differences between individuals and cognitive demands coming from the words to make the most of environmental opportunities and cognitive capacity.

In Experiment 3, we asked participants to read 1439 concrete and abstract nouns selected from a wide range of domains (e.g., temporal words such as “yesterday”, directional words such as “north”, emotional words such as “bewilderment” etc.). Then, participants were asked to associate words with locations on a precise, two-dimensional coordinate system. Thereby, we collected semantic coordinates for each word on the horizontal and vertical axes just as geographical coordinates in topographic environments in the real-world. For example, the word “moon” was located on the centre of the horizontal scale as of 17 pixels and on the upward side of the vertical scale as of 419.83 pixels. As such, “agony” was located on the leftward and downward position at (-243, -413.83). Location ratings showed that most of the words were positioned on a diagonal line from bottom left to top right. Words that denote objects were associated with locations where the objects typically occur in the real world. “Good” words with positive valence were associated with upward and rightward locations; whereas “bad” words with negative valence were associated with downward and leftward locations. Ratings revealed that vertical space was more dominant than horizontal space. That is, participants tended use the whole range of the vertical scale but rated the words more conservatively on the horizontal scale. Importantly, there was a high degree of agreement between participants in spatial ratings. We conclude that mappings between words and locations are even more consistent and fine-grained than previously documented (Dunn et al., 2014; Estes et al., 2015; Lachmair et al., 2011; Louwerse, 2008; Marmolejo-Ramos et al., 2013; Meteyard & Vigliocco, 2009). This is particularly interesting considering that language itself does not provide any rules or conventions as to spatial associations. In this respect, Experiment 3 results support a perceptual basis of language within grounded-embodied cognition (Barsalou, 1999; Pulvermüller, 1999).

Experiment 4 demonstrated that words, which are associated with locations in space either perceptually or metaphorically, give rise to simulations of these locations upon their presentation. We evidenced that language-based reactivations of space have consequences for memory in relation with physical locations and their simulations. Words that were presented in locations incongruent to the locations they imply (e.g., “moon” in a downward location or “agony” in an upward location) were remembered faster than words that were presented in congruent locations (e.g., “moon” in an upward location or “agony” in a downward location). A visual cue shown at retrieval guided participants’ attention to the previous locations of the probe word and/or semantically implied location. Performance was better when the cue location did not overlap with the location of the language-based spatial simulation. Consider the word “moon”. “Moon” guides attention to the upward side of the screen because it is associated with an upward location in space as Experiment 3 indicated. Experiment 4 showed that participants’ recognition memory was more accurate when they saw a cue in the bottom right quadrant, but it was slower when they saw a cue in the top left corner after encoding "moon".

Crucially, visual cues occurring in simulated locations deteriorated memory performance only with the modulation of imageability for words encoded in congruent positions. In other words, simulations invoked by high imageable words (but not low imageable words) interfered with visual cues if words were encoded in the congruent positions. Whereas, visual cues interfered with spatial simulations without any lexico-semantic modulation if the words were encoded in incongruent locations. Put differently, abstract words with spatial associations triggered strong spatial simulations and interfered with visual cues only when they were encoded in incongruent locations. We conclude that incongruency between physical and semantic space at the encoding stage produces an unexpectedness effect (Schmidt, 1991). Unexpectedness due to spatial incongruency has two impacts on memory for language: (1) It deepens memory traces which results in better retrieval performance. (2) It strengthens language-based spatial simulations which elicits a more robust interference effect.

Spatial simulations examined in this thesis have connections to language; yet, they are of two different types: (1) Simulation of locations (in which words are perceived) in the absence of words (offline spatial simulation). (2) Simulation of space suggested by word meanings (online spatial simulation). We evidenced that online and offline simulations of space have consequences on memory for language even if space does not seem to be a part of the retrieval process. We conclude that spatial perception and spatial cognition underlie memory and language operations in robust and systematic ways. In this respect, experiments provide compelling evidence for a memory and a language conceptualisation that is grounded in the sensorimotor system and extended into the body and the environment.

7.2 Implications

The overarching theme of the present thesis is space and spatial simulation. Experiments connect to each other in that space is involved in cognitive processing in different ways. Thus, the overall implication of the current work is showing that simulations and particularly, simulations of space play key roles in the human cognition. Along with that, experiments reported in this thesis extend into different domains of cognition; in particular, memory, language, grounded-embodied and extended cognition. Hence, results have considerable implications for the architecture of the abovementioned domains and potentially, on more practical fields.

7.2.1 Implications for Grounded-Embodied and Extended Cognition

Our results can be thought as further steps towards more precise models of grounded-embodied and extended cognition in a broader sense. In four experiments, we specified the mechanisms of simulation (Experiment 1, 2, 3 & 4) and cognitive offloading (Experiment 1 & 2). We observed a flexible and intelligent coordination between internal and external sources across all experiments. In Experiment 1 and 2, participants moved from internal to external bits of information wherever the "cost of an operation" (Kirsh, 2010) is lower according to their own sources and task conditions. In Experiment 4, the combination of external information (i.e., physical location) and internal information (i.e., semantic location) affected memory performance. In this respect, our results demonstrate not only the existence but also the importance of mental representations within grounded operations. Results support a relatively moderate view as compared to radical embodiment (Chemero, 2011) in terms of mental representations.

There were concrete and abstract words in all experiments. The effects (i.e., looking at nothing and language-induced spatial simulations) extended to both types of words. Hence, the current thesis suggests that abstract concepts can be grounded in space both perceptually (via their locations in the real-world) and semantically (via their meanings and metaphorical spatial associations). Grounding abstractness was not tested directly. Along with that, our findings seem to be more supportive for a hard approach to grounded-embodied cognition, which suggests that perceptual, motor and affective processes underlie all cognitive phenomena including those with abstractness (Barsalou, 1999; Pezzulo et al., 2013) rather than the pluralistic approach (Louwerse & Jeuniaux, 2010, 2012), which holds that abstract concepts are handled by symbolic systems.

Participants in Experiment 1 and 2 deployed memory work onto their body with eye movements and the environment with simulation of word locations. In this respect, we reframed looking at nothing as a type of cognitive offloading (Risko & Gilbert, 2016) and an efficient use of space and spatial simulation via cognitive offloading (Kirsh, 1995). Embedding looking at nothing into the abovementioned mechanisms and thus, extended cognition could conflate different but conceptually overlapping lines of research (i.e., looking at nothing and cognitive offloading) and open up novel directions in research and application (see Chapter 7.4).

In this respect, one potential implication could be making use of space and spatial simulation consciously in order to improve performance in tasks that require memory retrieval under demanding conditions. There are recommendations for active use of memory outsourcing with environmental modifications (e.g., external memory cues for navigating indoor environments such as hospitals and nursing homes etc.) in aging population (Ross & Schryer, 2015). As such, simulation of locations and looking at nothing might be operationalised in the service of memory in a strategic way. Individuals with memory disorders or learning disabilities, older population, students, learners of second language or professionals in memory-demanding jobs (e.g., interpreters, air traffic controllers etc.) (G. G. Fisher et al., 2014) can be trained to use the environment effectively to support their internal memory.

7.2.2 Implications for Memory

Results from Experiment 1, 2 and 4 extend the contemporary memory research in agreement with theories of grounded-embodied and extended cognition. There are two crucial tenets of the memory model that the experiments reported in the current thesis present:

(1) Remembering is a partial simulation of encoding: During retrieval, participants reactivated their spatio-perceptual experiences they had at encoding (i.e., perceiving word locations on the screen). We conclude that remembering words involves reconstruction of the perceptual states that were present during encoding (Danker & Anderson, 2010). In short, remembering can be thought as a partial simulation of encoding. Simulation of encodings is "partial" rather than an exact "replay" because simulations themselves are "sketchy" and "never complete" (Barsalou, 1999). Thus, we focused on the percentage of looks in the relevant locations rather than the correspondence of eye movements executed at encoding and at retrieval (cf., Laeng & Teodorescu, 2002) (see also Chapter 7.4.1). Findings fit well with the experimental evidence that encoding and retrieval are strongly interdependent and common neural systems are activated in both processes (Goldberg et al., 2006b; Nyberg et al., 2000; Otten, 2007; Rugg et al., 2008; Wheeler et al., 2000).

(2) Remembering involves perceptual and contextual details of encoding: Retrieval of words involve spatial locations (Experiment 1 and 2) and semantically suggested locations (Experiment 4). Further, memory for language was affected by spatial perception although participants in these experiments were not aware of the link between the words and the spatial manipulation. Results underline the perceptual underpinnings of memory within grounded-embodied cognition. Simulation-based memory model in grounded-embodied cognition and thus, results in this thesis contrast to computational models (e.g., multi-store model in Atkinson and Shiffrin (1968)), which hold that memory and sensory systems are detached from each other (Buckner & Wheeler, 2001).

As a novel contribution, our findings tap into a critical question as to the simulations in memory: How does accessibility of retrieved information affect simulation of encoding? Danker and Anderson (2010) highlighted this question as an important research avenue among others. Results in the present thesis demonstrated that items that are more difficult to access result in more robust activations through eye movements. We conclude that memory for language is a flexible process of piecing together fragments of both semantic and perceptual information and it is based on partial reconstructions of the sensorimotor experiences of encoding. Simulation approach to memory could have even broader implications. First, understanding the processes and neural substrates involved in visual and auditory perception, for instance, could shed light on the role of these processes and substrates in remembering words (Buckner & Wheeler, 2001). Second, investigating perceptual and thus, encoding dysfunctions could clarify memory distortions among healthy and clinical population (e.g., in amnesia, dementia or schizophrenia) (Michaelian, 2016a; Schacter, Guerin, & St. Jacques, 2011; Tek et al., 2002).

7.2.3 Implications for Language

Language operations in the present thesis resulted in spatio-perceptual simulations even though space was never an explicit part of the task with the exception of Experiment 3. In this respect, results (particularly those from Experiment 4) extend the literature of simulation semantics (Bergen, 2007). The simulation approach to language and thus, our findings have implications for the general architecture of language. Computational models of language argue that language is a system of abstract symbols that reside in an amodal region of the brain independent of the perceptual and the motor system and that language comprehension is a manipulation of these abstract symbols (e.g., Chomsky, 1980; May, 1985; Pinker, 1995). Our findings, along with the overwhelming evidence that language comprehension involves the activation of sensorimotor experiences (see Chapter 2.1.3), disapprove the computational view.

The novel implication of our results is the specification of the linguistic properties that give rise to spatial simulations. First, our findings demonstrate that lexical items are sufficient to generate simulations without any contextual information or without integration into larger linguistic structures such as sentences. This finding is in contrast with the results in Bergen, Lindsay, Matlock and Narayanan (2007), who demonstrated that comprehension of the sentence as a whole and not simply lexical associations yields simulations. Second, our findings demonstrate that word imageability is a crucial property both for looking at nothing and for the subsequent effects of language-induced simulations. We show that less imageable/more abstract words lead to more frequent reliance on simulation of word locations (Experiment 2). On the other hand, imageability plays an important role in exacerbating the word-induced simulations when the physical arrangement dampens their effects (Experiment 4). Third, our findings demonstrate that abstract words are grounded in space as well as concrete words and they trigger robust simulations which later impact performance. As a matter of fact, Experiment 3 results indicate that emotional abstract words had even more spatiality (as opposed to being spatially neutral) than concrete words with lower emotionality. Grounding of abstract language in space is informative not only for theories of grounded-embodied cognition but also for theories of language processing considering the previous evidence that metaphorical language does not yield simulations (Bergen et al., 2007) or lead to different simulation effects as compared to concrete language (Liu & Bergen, 2016).

In conclusion, the present thesis puts forward that language is grounded in space and spatial perception in a more robust manner than previously documented. Spatial simulations triggered by language have consequences on tasks involving language. Therefore, language models and experimental studies should consider spatial domain even when linguistic stimuli in question do not explicitly denote spatial positions or when spatial perception or spatial simulations are not investigated. Considering the fundamentality of space in language and robustness of language-induced simulations, we call this domain *spatial-simulation semantics* combining the terms *spatial semantics* (Zlatev, 2012) and simulation semantics (Bergen, 2007).

7.3 Future Work

The research on looking at nothing, cognitive offloading and simulation mechanism within the framework of grounded-embodied and extended cognition is still in its infancy. Hence, there are numerous outstanding questions waiting to be answered. In general, looking at nothing via simulation of locations and language-based spatial simulations should be further specified as detailed below. The questions and issues listed below can direct a research path towards well-defined models of both mechanisms.

7.3.1 Looking at nothing

1. At what point exactly do individuals look at blank locations? We show that fixations in the relevant, blank locations increase as memory load increases as a function of word difficulty in Experiment 2. More studies are required to specify the effect of memory load on looking at nothing. For example, future studies could measure mental load based on ocular indices such as pupil dilation (van der Wel & van Steenbergen, 2018)

and ocular aberrations (i.e., imperfections of the ocular optic elements arranged along the ocular axis) (Jiménez, Cárdenas, González-Anera, Jiménez, & Vera, 2018) or other biological markers such as heart rate variability (Cerpa, Chandler, & Sweller, 1996), facial thermography (Marinescu et al., 2018) or galvanic skin response (Kohlisch & Schaefer, 1996). Based on more precise measures, it might be possible to identify the amount of cognitive load that is required to guide participants to environmental sources.

2. How precisely are word locations registered? Consider that participants are asked to retrieve specific words in a sentence or in a longer text. In such a case, do fixations fall specifically on the previous word locations or within their periphery in a sentence, or a larger text? If, for example, fixations accumulate in a larger area corresponding to the word to be retrieved rather than the word's specific location, this could be informative about the spatial precision of indexing and accordingly, requirements of looking at nothing. Related to this, how does spatial precision relate to functionality of looking at nothing? There could be a positive correlation between precision in spatial indexing and functionality of fixations in looking at nothing. Future studies could also make use of location ratings collected in Experiment 3 to investigate the effect of spatial position on looking behaviour.

3. How much information can be attached to one location? Experiment 1 and 2 demonstrate that subsequent visual cues update spatial indices corresponding to words (see also Richardson & Kirkham, 2004). However, there is no direct evidence that more than one word can be registered to the same location. In such a case, do individuals look at the previous location equally often for each word; or does the frequency of looks change as a function of the encoding order?

4. Future studies should further clarify the functionality of eye movements in memory. The studies evidencing that looks to blank locations improve memory performance (Johansson & Johansson, 2014; Scholz et al., 2018, 2016) follow a similar methodology. It is critical to settle whether looks to blank locations benefit memory at all times; or whether specific conditions (e.g., memory load at a certain threshold, spatial distance between the visual cue at retrieval and the previous word location etc.) should be met. Also, new methodologies should be developed, in which participants are not forced to look at certain locations during retrieval, with the aim of gathering more reliable and ecologically valid data (see Wantz et al., 2016 as an example).

5. What is the time course of looking at nothing? We observed that the majority of the participants looked at previous locations of the words at the very beginning and the very end of the retrieval phase. Are there such specific time windows within the retrieval phase during which looking at nothing increases and/or decreases? Are looking at nothing patterns within these time windows comparable to each other or do they have different characteristics (e.g., functional eye movements at early stages and confirmatory looks at later stages)? As an example, Martarelli, Mast and Hartmann (2017) performed a time-course analysis on the spontaneous eye movements during free recall and recognition of past and future items and showed that position changes (e.g., rightward looks for future items) took place within the first 50% of the recall and recognition duration. A similar analysis can be applied to our looking at nothing paradigm.

6. How does the looking behaviour during encoding relate to the looking behaviour during retrieval? Grounded-embodied theories of memory emphasise the correspondence between encoding and retrieval (Kent & Lamberts, 2008). In this regard, for example,

are there similarities between the time course of eye movements at encoding and retrieval and does such a potential overlap predict the functionality of eye movements (Laeng & Teodorescu, 2002; Noton & Stark, 1971)? There is evidence that eye movements at retrieval are not exact reinstatements of the movements occurring at encoding for complex images (Johansson et al., 2012). However, there is no evidence if this is the case for visually presented single words.

7. What is the neural basis of looking at nothing? We speculate that encoding words on a grid taps into the regions of brain that are specialised in processing spatial information. This hypothesis seems probable considering the relation between visuospatial memory and looking at nothing behaviour. However, we do not have direct evidence to support this hypothesis to our knowledge. Further, event-related brain potentials could provide a new perspective to the time course of looking behaviour during encoding and retrieval. Potentially different activations (e.g., N400) at consecutive time windows (e.g., early vs. late stages) within the retrieval phase might illuminate the characteristics of looking behaviour.

8. Do looks to blank locations always occur without awareness? Future studies should be designed in a way to investigate whether looking at nothing is a completely unconscious behaviour, or whether we have some kind of control on our “decision” to offload memory work onto the world (Risko & Gilbert, 2016). Informal queries with the participants following Experiment 1 and 2 indicated that participants were unaware of the spatial arrangement and that they looked at blank locations without awareness. Future studies could make use of subjective or objective measures of consciousness based on behavioural and/or neural methods (Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008). For example, participants can be formally asked to make metacognitive

judgements about their memory “strategies” during the task. If looking at nothing is an entirely unconscious behaviour, participants would not know when they rely on the environment and when they rely on internal sources. Pupillary responses can support the investigation of consciousness in looking at nothing because pupil diameters continuously change in response to task conditions without awareness (Laeng et al., 2012). Event-related brain potentials and fMRI can be employed for the same aim as consciously perceived stimuli (simulation of previous word locations in our case) elicit widespread brain activation and trigger different ERP signals as compared to stimuli that do not reach consciousness (Baars, 2002; Seth et al., 2008). On the other hand, if looking at nothing is a partly conscious behaviour, it is important to identify which factors (e.g., metacognitive believes and confidence in internal sources) modulate decisions to rely on the environment. This question can be investigated in a study where looking at nothing is given as a deliberate memory strategy and participants are instructed in this way (Risko & Dunn, 2015).

9. Following on from question 8 above: Can looking at nothing be taught and/or improved? If so, is it possible “motivate” individuals with better visuospatial memory and/or worse mental imagery ability to rely more on the environmental sources? Does such a memory training translate into a meaningful improvement in memory performance?

10. Can spatial indexing be prevented or controlled to a certain degree? Our findings indicate that individuals index word locations automatically. However, an experimental paradigm manipulating the visual/contextual information on the screen with masking or filtering could identify the visual properties that enable or contribute to spatial indexing of words.

7.3.2 Language-based simulations of space

1. We introduced a variable of spatial-simulation semantics that we call 'spatiality' in Experiment 3. Spatiality shows the distance from spatially neutral (i.e., located in the centre of the rating screen) for a given word. A similar variable, that is, spatial iconicity, refers to the strength of the mapping between the word and its location (e.g., Louwerse, 2008). We predict that these two concepts refer to different aspects of language-space associations. That is, a word located in the centre of the rating screen (thus, spatially neutral) could have a high spatial iconicity (strongly associated with central location) such as the word "centre" and vice versa. Norming studies should test the correlation between spatiality and spatial iconicity to confirm this prediction. We speculate that words with higher spatial iconicity could lead to stronger simulations of space.

2. Future studies could look at the relation between handedness and associations between words and spatial locations. There is evidence that individuals associate words with positive valence with the side of space that corresponds to their dominant hand and words with negative valence with the side of their nondominant hand (i.e., body-specifity hypothesis) (Casasanto, 2009; de la Vega et al., 2012). It could be important to verify body-specificity hypothesis when participants are explicitly asked to rate word locations in a norming study (pure association) rather than attending a task with an implicit spatial component. Further, cross-linguistic norming studies are necessary to clarify whether languages differ in the mappings between words and spatial locations.

3. Future studies with event-related brain potentials could provide valuable information on the effect of spatial incongruency between word location and location implied by word meaning. Particularly, we predict a potential modulation of the N400 component

during the encoding of spatially incongruent words. Because N400 is sensitive to perceptual modality switching (Hald et al., 2011) and semantic violations (e.g., implausible adjective-noun combinations) (Hagoort, 2003). Event-related potentials could also give invaluable information about the timing of events within the encoding stage as opposed to end-state variables of memory performance (hit rate and hit latency) (Kutas & Federmeier, 2011).

4. How does conscious mental imagery ability map onto language-based spatial simulations? Experiment 2 results show that both word imageability and self-reports of visual imagery are correlated with the tendency to rely on the environment via spatial simulations. However, there is also evidence showing that there is no systematic relation between vividness of visual imagery (based on questionnaires) and the amplitude of modality-switching effect based on simulations (Pecher, van Dantzig, & Schifferstien, 2009). A more refined version of Experiment 4 with objective measures of mental imagery could examine whether participants with better imagery abilities are subject to greater mismatch advantage and interference effects due to language-based spatial simulations.

5. Can language-induced simulations be prevented? Our findings indicate that simulations are automatic and cannot be controlled. If it is possible to stop the simulations activated by words, then their effects could be modulated. An experiment aiming at this could be informative in understanding the automaticity of simulations activated by language.

List of References

Aguiar, A., & Baillargeon, R. (2002). Developments in young infants' reasoning about occluded objects. *Cognitive Psychology*, *45*(2), 267–336. https://doi.org/10.1016/S0010-0285(02)00005-1

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Alibali, M. W., & Dirusso, A. A. (1999). The function of gesture in learning to count: more than keeping track. *Cognitive Development*, *14*(1), 37–56. https://doi.org/10.1016/S0885-2014(99)80017-3

Allan, K., Robb, W. G. K., & Rugg, M. D. (2000). The effect of encoding manipulations on neural correlates of episodic retrieval. *Neuropsychologia*, *38*(8), 1188–1205. https://doi.org/10.1016/S0028-3932(00)00013-0

Althoff, R. R., & Cohen, N. J. (1999). Eye movement based memory effect: a reprocessing effect in face perception. *Journal of Experimental Psychology*, *25*(4), 997–1010. https://doi.org/10.1037/0278-7393.25.4.997

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, *93*(2), B79-87. https://doi.org/10.1016/j.cognition.2004.02.005

Amorim, M., & Pinheiro, A. P. (2018). Is the sunny side up and the dark side down? Effects of stimulus type and valence on a spatial detection task. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2018.1452718

Andrade, J., & Meudell, P. (1993). Short report: Is spatial information encoded automatically in memory? *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *46*(2), 365–375. https://doi.org/10.1080/14640749308401051

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70. https://doi.org/10.1037/h0036799

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In K. W. Spence & J. T. B. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). Oxford: Academic Press. https://doi.org/10.1016/s0079-7421(08)60422-3

Avraamides, M. N. (2003). Spatial updating of environments described in texts. *Cognitive Psychology*, *47*(4), 402–431. https://doi.org/10.1016/S0010-0285(03)00098-7

Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database. Philadelphia, PA.: Linguistic Data Consortium, University of Pennsylvania.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *The Behavioral and Brain Sciences*, *20*(4), 723-742; discussion 743-767. https://doi.org/10.1017/S0140525X97001611

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577-609; discussion 610-660. https://doi.org/10.1017/S0140525X99252144

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*(5–6), 513–562. https://doi.org/10.1080/01690960344000026

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2010). Grounded cognition: past, present, and future. *Topics in Cognitive Science*, *2*(4), 716–724. https://doi.org/10.1111/j.1756-8765.2010.01115.x

Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergen, B. K. (2005). Mental simulation in spatial language processing. In B. G. B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 232–237). Austin, TX: Cognitive Science Society.

Bergen, B. (2007). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 277-301). Amsterdam: John Benjamins. http://doi.org/10.1075/hcp.18.19ber

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*(5), 733–764. https://doi.org/10.1080/03640210701530748

Blignaut, P., & Wium, D. (2014). Eye-tracking data quality as affected by ethnicity and experimental design. *Behavior Research Methods*, *46*(1), 67–80. https://doi.org/10.3758/s13428-013-0343-0

Bloom, P., Garrett, M. F., Nadel, L., & Peterson, M. A. (Eds.). (1996). *Language and space*. Cambridge, MA: The MIT Press.

Bochynska, A., & Laeng, B. (2015). Tracking down the path of memory: Eye scanpaths facilitate retrieval of visuospatial information. *Cognitive Processing*, *16*(1), 159–163.

https://doi.org/10.1007/s10339-015-0690-0

Boot, I., & Pecher, D. (2011). Representation of categories: metaphorical use of the container schema. *Experimental Psychology*, *58*(2), 162–170. https://doi.org/10.1027/1618-3169/a000082

Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28. https://doi.org/10.1016/S0010-0277(99)00073-6

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, *13*(2), 185–189. https://doi.org/10.1111/1467-9280.00434

Borreggine, K. L., & Kaschak, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science*, *30*(6), 1097–1112. https://doi.org/10.1207/s15516709cog0000_91

Borst, G., & Kosslyn, S. M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory and Cognition*, *36*(4), 849–862. https://doi.org/10.3758/MC.36.4.849

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*(1), 27–38. https://doi.org/10.1162/jocn.1997.9.1.27

Brewer, J. B., Zhao, Z., Glover, G. H., & Gabrieli, J. D. E. (1998). Making memories: Brain activity that predicts whether visual experiences will be remembered or forgotten. *Science*, *281*(5380), 1185–1187. https://doi.org/10.1126/science.281.5380.1185

Bridgeman, B., Van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*(2), 247–258. https://doi.org/https://doi.org/10.1017/S0140525X00034361

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–11. https://doi.org/10.3758/s13428-013-0403-5

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research*, *24*(3), 355–363. https://doi.org/10.1016/j.cogbrainres.2005.02.020

Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nature Reviews Neuroscience*, *2*(9), 624–634. https://doi.org/10.1038/35090048

Buswell, G. T. (1936). How people look at pictures. *Psychological Bulletin*, *33*(2), 142–143. https://doi.org/10.1037/h0053409

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*(1), 3–18. https://doi.org/10.1111/j.1469-7610.2004.00374.x

Caggiano, V., Fogassi, L., Rizzolatti, G., Casile, A., Giese, M. A., Thier, P., … Rizzolatti, G. (1996). Premotor cortex and the recognition of motor actions. *Current Biology*, *23*(2), 131–141. https://doi.org/10.1016/0926-6410(95)00038-0

Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, *37*(3), 411–437. https://doi.org/10.1006/jmla.1997.2519

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1999). Time-course of fMRI activation in langauge and spatial networks during sentence comprehension. *NeuroImage*, *10*(2), 216–224. https://doi.org/https://doi.org/10.1006/nimg.1999.0465

Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right- and left-handers. *Journal of Experimental Psychology. General*, *138*(3), 351–367. https://doi.org/10.1037/a0015854

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

Casasanto, D., & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition*, *115*(1), 179–185. https://doi.org/10.1016/j.cognition.2009.11.002

Casasola, M. (2005). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology*, *41*(1), 183–192. https://doi.org/10.1037/0012-1649.41.1.183

Casasola, M. (2008). The development of infants' spatial categories. *Current Directions in Psychological Science*, *17*(1), 21–25. https://doi.org/10.1111/j.1467-8721.2008.00541.x

Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., & Pietrini, P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience and Biobehavioral Reviews*, *32*(8), 1346–1360. https://doi.org/10.1016/j.neubiorev.2008.05.002

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research*, *15*(4), 345–367. https://doi.org/10.2190/MG7X-4J8N-CKYR-P06T

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, *12*(4), 478–484. https://doi.org/10.1006/nimg.2000.0635

Chasteen, A. L., Burdzy, D. C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia*, *48*(2), 627–630. https://doi.org/10.1016/j.neuropsychologia.2009.09.029

Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge, MA: MIT press.

Chomsky, N. A. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(1), 1–15. https://doi.org/doi:10.1017/S0140525X00001515

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences*, *21*(2), 111–124. https://doi.org/10.1016/j.tics.2016.12.007

Cicchetti, D. V. (1994). Guidlines, criteria, and rules of thumb for evalauting normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cichy, R. M., Heinzle, J., & Haynes, J. D. (2012). Imagery and perception share cortical representations of content and location. *Cerebral Cortex*, *22*(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Clark, A. (1989). *Microcognition*. Cambridge, MA: MIT Press.

Clark, A. (1998). *Being there: putting brain, body, and world together again*. Cambridge, MA: MIT press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1111/j.1467-9744.2009.01021.x

Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, and Computers*, *36*(3), 371–383. https://doi.org/10.3758/BF03195584

Coll-Florit, M., & Gennari, S. P. (2011). Time in language: event duration in language comprehension. *Cognitive Psychology*, *62*(1), 41–79. https://doi.org/10.1016/j.cogpsych.2010.09.002

Collette, F., Majerus, S., Van der Linden, M., Dabe, P., Degueldre, C., Delfiore, G., … Salmon, E. (2001). Contribution of lexico-semantic processes to verbal short-term memory tasks: a PET activation study. *Memory*, *9*(4), 249–259. https://doi.org/10.1080/09658210143000056

Coltheart, V., & Winograd, E. (1986). Word imagery but not age of acquisition affects episodic memory. *Memory & Cognition*, *4*(2), 174–180. https://doi.org/10.3758/BF03198377

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*(3), 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

Conway, A. R., & Engle, R. W. (1996). Individual differences in working memory capacity: more evidence for a general capacity theory. *Memory*, *4*(6), 577–590. https://doi.org/10.1080/741940997

Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, *26*(2), 199–208. https://doi.org/10.1017/S0140525X03000062

Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Experimental Brain Research*, *192*(3), 553–560. https://doi.org/10.1007/s00221-008-1491-9

Corsi, P. (1972). *Human memory and the medial temporal lobe region of the brain*. Doctoral thesis. Montreal: McGill University.

Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2,578 monosyllabic words. *Memory*, *18*(6), 595–609. https://doi.org/10.1080/09658211.2010.493892

Cortese, M. J., McCarty, D. P., & Schock, J. (2015). A mega recognition memory study of 2897 disyllabic words. *Quarterly Journal of Experimental Psychology*, *68*(8), 1489–1501. https://doi.org/10.1080/17470218.2014.945096

Coslett, H. B. (1999). Spatial influences on motor and language function. *Neuropsychologia*, *37*(6), 695–706. https://doi.org/10.1016/S0028-3932(98)00116-X

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Craver-Lemley, C., & Reeves, A. (1992). How visual imagery interferes with vision. *Psychological Review*, *99*(4), 633–649. https://doi.org/http://dx.doi.org/10.1037/0033-295X.99.4.633

Damjanovic, L., & Santiago, J. (2016). Contrasting vertical and horizontal representations of affect in emotional visual search. *Psychonomic Bulletin & Review*, *23*(1), 62–73. https://doi.org/10.3758/s13423-015-0884-6

Danker, J. F., & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin*, *136*(1), 87–102. https://doi.org/10.1037/a0017937

De Brigard, F. (2014). The nature of memory traces. *Philosophy Compass*, *9*(6), 402–414. https://doi.org/10.1111/phc3.12133

de Groot, F., Huettig, F., & Olivers, C. N. L. (2016). Revisiting the looking at nothing phenomenon: visual and semantic biases in memory search. *Visual Cognition*, *24*(3), 226–245. https://doi.org/10.1080/13506285.2016.1221013

de Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse Processes*, *54*(7), 493–503. https://doi.org/10.1080/0163853X.2015.1119604

de la Fuente, J., Casasanto, D., Roman, A., & Santiago, J. (2015). Can culture influence body-specific associations between space and valence? *Cognitive Science*, *39*(4), 821–832. https://doi.org/10.1111/cogs.12177

de la Vega, I., de Filippis, M., Lachmair, M., Dudschig, C., & Kaup, B. (2012). Emotional valence and physical space: limits of interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 375–385. https://doi.org/10.1037/a0024979

De la Vega, I., Dudschig, C., De Filippis, M., Lachmair, M., & Kaup, B. (2013). Keep your hands crossed: the valence-by-left/right interaction is related to hand, not side, in an incongruent hand-response key assignment. *Acta Psychologica*, *142*(2), 273–277. https://doi.org/10.1016/j.actpsy.2012.12.011

De Renzi, E., & Spinnler, H. (1967). Impaired performance on color tasks in patients with hemispheric damage. *Cortex*, *3*(2), 194–217. https://doi.org/https://doi.org/10.1016/S0010-9452(67)80012-1

de Vega, M., León, I., Hernández, J. A., Valdés, M., Padrón, I., & Ferstl, E. C. (2014). Action sentences activate sensory motor regions in the brain independently of their status of reality. *Journal of Cognitive Neuroscience*, *26*(7), 1363–1376. https://doi.org/10.1162/jocn_a_00559

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396.

https://doi.org/10.1037/0096-3445.122.3.371

Dewhurst, S. A., Hitch, G., & Barry, C. (1998). Separate effects of word frequency and age of acquisition in recogntion and recall. *Journal of Experimental Psychology*, *24*(2), 284–298. https://doi.org/10.1037/0278-7393.24.2.284

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition*, *102*(1), 139–149. https://doi.org/10.1016/j.cognition.2005.12.009

Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers in Psychology*, *6*(October). https://doi.org/10.3389/fpsyg.2015.01525

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: a response to cognitive or social difficulty? *Memory & Cognition*, *33*(4), 727–733. https://doi.org/10.3758/BF03195338

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "sun" and looking up: the influence of language on saccadic eye movements in the vertical dimension. *PLoS ONE*, *8*(2), e56872. https://doi.org/10.1371/journal.pone.0056872

Dunn, B., Kamide, Y., & Scheepers, C. (2014). Hearing "moon" and looking up: Word-related spatial associations facilitate saccades to congruent locations. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 433–438). Austin, TX: Cognitive Science Society. https://doi.org/10.13140/2.1.2952.3529

Ehrlichman, H., & Barrett, J. (1983). "Random" saccadic eye movements during verbal-linguistic and visual-imaginal tasks. *Acta Psychologica*, *53*(1), 9–26. https://doi.org/10.1016/0001-6918(83)90013-6

Ehrlichman, H., & Micic, D. (2012). Why do people move their eyes when they think? *Current Directions in Psychological Science*, *21*(2), 96–100. https://doi.org/10.1177/0963721412436810

Ellis, N. R. (1991). Automatic and effortful processes in memory for spatial location. *Bulletin of the Psychonomic Society*, *29*(1), 28–30. https://doi.org/10.3758/BF03334759

Emmorey, K. (2001). Space on hand: The exploitation of signing space to illustrate abstract thought. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 147–174). Cambridge, MA: MIT Press.

Emmorey, K., Tversky, B., & Taylor, H. A. (2000). Using space to describe space: perspective in speech, sign, and gesture. *Spatial Cognition and Computation*, *2*(3), 157–180. https://doi.org/10.1023/A:1013118114571

Estes, Z., Verges, M., & Adelman, J. S. (2015). Words, objects, and locations: Perceptual matching explains spatial interference and facilitation. *Journal of Memory and Language*, *84*, 167–189. https://doi.org/10.1016/j.jml.2015.06.002

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, *19*(2), 93–97. https://doi.org/10.1111/j.1467-9280.2008.02051.x

Eysenck, M. W., & Eysenck, M. C. (1980). Effects of processing depth, distinctiveness, and word frequency on retention. *British Journal of Psychology*, *71*(2), 263–274.

https://doi.org/10.1111/j.2044-8295.1980.tb01743.x

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, *15*(2), 399–402. https://doi.org/10.1046/j.0953-816x.2001.01874.x

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *12*(11), 405–410. https://doi.org/10.1016/j.tics.2008.07.007

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, *87*(2), 113–123. https://doi.org/10.1037/h0021465

Fischer, M. H. (1999). Memory for word locations in reading. *Memory*, *7*(1), 79–116. https://doi.org/10.1080/741943718

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555–556. https://doi.org/10.1038/nn1066

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, *61*(6), 825–850. https://doi.org/10.1080/17470210701623605

Fiser, A., Mahringer, D., Oyibo, H. K., Petersen, A. V., Leinweber, M., & Keller, G. B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nature Neuroscience*, *19*(12), 1658–1664. https://doi.org/10.1038/nn.4385

Fisher, G. G., Infurna, F. J., Grosch, J., Stachowski, A., Faul, J. D., & Tetrick, L. E. (2014). Mental work demands, retirement, and longitudinal trajectories of cognitive functioning. *Journal of Occupational Health Psychology*, *19*(2), 231–242. https://doi.org/10.1037/a0035724

Fisher, R. P., & Craik, F. I. M. (1980). The effects of elaboration on recognition memory. *Memory & Cognition*, *8*(5), 400–404. https://doi.org/10.3758/BF03211136

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage*, *32*(3), 1413–1421. https://doi.org/10.1016/j.neuroimage.2006.06.007

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Foulsham, T., & Kingstone, A. (2012). Fixation-dependent memory for natural scenes: an experimental test of scanpath theory. *Journal of Experimental Psychology: General*, *142*(1), 41–56. https://doi.org/10.1037/a0028227

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2), 6.1-17. https://doi.org/10.1167/8.2.6

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, *119*(1), 63–76. https://doi.org/10.1037/0096-3445.119.1.63

Frick, A., & Möhring, W. (2013). Mental object rotation and motor development in 8- and 10-month-old infants. *Journal of Experimental Child Psychology*, *115*(4), 708–720.

https://doi.org/10.1016/j.jecp.2013.04.001

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*(4), 375–399. https://doi.org/10.3758/BF03203275

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*, 87–115. https://doi.org/10.1146/annurev.psych.49.1.87

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. https://doi.org/10.1093/brain/119.2.593

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5

Gallese, V., & Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3–4), 455–479. https://doi.org/10.1080/02643290442000310

Galli, G. (2014). What makes deeply encoded items memorable? Insights into the levels of processing framework from neuroimaging and neuromodulation. *Frontiers in Psychiatry*, *5*(May), 61. https://doi.org/10.3389/fpsyt.2014.00061

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, *20*(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Garnham, A. (1981). Mental models as representations of text. *Memory & Cognition*, *9*(6), 560–565. https://doi.org/https://doi.org/10.3758/BF03202350

Gasser, M. (2004). The origins of arbitrariness in language. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 434–439). Austin, TX: Cognitive Science Society. https://doi.org/10.1098/rstb.2013.0299

Gibbs, R. W. (1994). *The poetics of mind: figurative thought, language, and understanding*. Cambridge: Cambridge University Press.

Gibbs, R. W. (2005). *Embodiment and cognitive science*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511805844

Gibson, J. J. (1979). *The ecological approach to visual perception*. Hove: Psychology Press.

Giessner, S. R., & Schubert, T. W. (2007). High in the hierarchy: How vertical location and judgments of leaders' power are interrelated. *Organizational Behavior and Human Decision Processes*, *104*(1), 30–44. https://doi.org/10.1016/j.obhdp.2006.10.001

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, *68*(5), 971–992. https://doi.org/10.1080/17470218.2014.972963

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(1), 21–31. https://doi.org/10.1037/0278-7393.2.1.21

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–19. https://doi.org/10.1017/S0140525X97000010

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*(3), 558–565. https://doi.org/10.3758/BF03196313

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83. https://doi.org/10.1016/0749-596X(87)90063-5

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition*, *26*(4), 651–658. https://doi.org/10.3758/BF03211385

Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience*, *26*(1), 133–179. https://doi.org/10.1146/annurev.neuro.26.010302.081134

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006a). Distinct and common cortical activations for multimodal semantic categories. *Cognitive, Affective and Behavioral Neuroscience*, *6*(3), 214–222. https://doi.org/10.3758/CABN.6.3.214

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006b). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience*, *26*(18), 4917–4921. https://doi.org/10.1523/JNEUROSCI.5389-05.2006

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, *3*(11), 419–429. https://doi.org/10.1016/S1364-6613(99)01397-2

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences*, *9*(5), 234–241. https://doi.org/10.1016/j.tics.2005.03.006

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, *21*(2), 90–95. https://doi.org/10.1177/0963721412436811

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *NeuroImage*, *32*(2), 906–912. https://doi.org/10.1016/j.neuroimage.2006.03.037

Goodhew, S. C., McGaw, B., & Kidd, E. (2014). Why is the sunny side always up? Explaining the spatial mapping of concepts by language use. *Psychonomic Bulletin & Review*, *21*(5), 1287–1293. https://doi.org/10.3758/s13423-014-0593-6

Gordon, R. (1949). An investigation into some of the factors that favour the formation of stereotyped images. *British Journal of Psychology. General Section*, *39*(3), 156–167. https://doi.org/10.1111/j.2044-8295.1949.tb00215.x

Gozli, D. G., Chasteen, A. L., & Pratt, J. (2012). The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology: General*, *142*(4), 1028–1046. https://doi.org/10.1037/a0030362

Gozli, D. G., Pratt, J., Martin, K. Z., & Chasteen, A. L. (2016). Implied spatial meaning and visuospatial bias: Conceptual processing influences processing of visual targets and distractors. *PLoS ONE*, *11*(3), e0150928. https://doi.org/10.1371/journal.pone.0150928

Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., … Haxby, J. V. (1995). Age-related reductions in human recognition memory due to impaired encoding. *Science*, *269*(5221), 218–221. https://doi.org/10.1126/science.7618082

Gunraj, D. N., Upadhyay, S. S. N., Houghton, K. J., Westerman, D. L., & Klin, C. M. (2017). Simulating a story character's thoughts: evidence from the directed forgetting task. *Journal of Memory and Language*, *96*, 1–8. https://doi.org/10.1016/j.jml.2017.04.005

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899. https://doi.org/10.1162/089892903322370807

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (Vol. 7). London: Pearson.

Hald, L. A., Marshall, J.-A., Janssen, D. P., & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology*, *2*(45). https://doi.org/http://doi.org/10.3389/fpsyg.2011.00045

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. https://doi.org/10.20982/tqmp.08.1.p023

Hanakawa, T. (2016). Organizing motor imageries. *Neuroscience Research*, *104*, 56–63. https://doi.org/10.1016/j.neures.2015.11.003

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*(October), 166. https://doi.org/10.3389/fnhum.2010.00166

Hannula, D. E., & Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron*, *63*(5), 592–599. https://doi.org/10.1016/j.neuron.2009.08.025

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences*, *95*(25), 15061–15065. https://doi.org/10.1073/pnas.95.25.15061

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1–3), 335–346. https://doi.org/10.1016/0167-2789(90)90087-6

Hartmann, M., Martarelli, C. S., Mast, F. W., & Stocker, K. (2014). Eye movements during mental time travel follow a diagonal line. *Consciousness and Cognition*, *30*, 201–9. https://doi.org/10.1016/j.concog.2014.09.007

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356–388. https://doi.org/10.1037/0096-3445.108.3.356

Haugeland, J. (1985). *Artificial intelligence: the very idea*. Cambridge, MA: MIT Press.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307. https://doi.org/10.1016/S0896-6273(03)00838-9

Hauk, O., & Tschentscher, N. (2013). The body of evidence: What can neuroscience tell us about embodied semantics. *Frontiers in Psychology*, *4*(February), 50. https://doi.org/10.3389/fpsyg.2013.00050

Haun, D. B. M., Rapold, C. J., Janzen, G., & Levinson, S. C. (2011). Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition*, *119*(1), 70–80. https://doi.org/10.1016/j.cognition.2010.12.009

Hayhoe, M. M., & Ballard, D. H. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research*, *38*(1), 125–137. https://doi.org/10.1016/S0042-6989(97)00116-8

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, *55*(1), 39–84. https://doi.org/10.1016/0010-0277(94)00643-Y

Heaver, B., & Hutton, S. B. (2011). Keeping an eye on the truth? Pupil size changes associated with recognition memory. *Memory (Hove, England)*, *19*(4), 398–405. https://doi.org/10.1080/09658211.2011.575788

Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, *75*(6), 466–477. https://doi.org/10.1037/h0026771

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–285. https://doi.org/10.1016/j.tics.2004.04.001

Heisz, J. J., & Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, *8*(1), 1–10. https://doi.org/10.1167/8.1.9.Introduction

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: detecting changes to saccade targets in scenes. *Perception & Psychophysics*, *65*(1), 58–71. https://doi.org/10.3758/bf03194783

Henderson, J. M., Shinkareva, S. V., Wang, J., Luke, S. G., & Olejarczyk, J. (2013). Predicting cognitive state from eye movements. *PLoS ONE*, *8*(5), e64937. https://doi.org/10.1371/journal.pone.0064937

Hespos, S. J., & Rochat, P. (1997). Dynamic mental representation in infancy. *Cognition*, *64*(2), 153–188. https://doi.org/https://doi.org/10.1016/S0010-0277(97)00029-2

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature*, *1*, 453–456. https://doi.org/10.1038/nature02634

Hesslow, G. (2011). The current status of the simulation theory of cognition. *Brain Research*, *1428*, 71–79. https://doi.org/10.1016/j.brainres.2011.06.026

Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews*, *80*(3), 953–978. https://doi.org/http://physrev.physiology.org/content/80/3/953

Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). Cambridge: Cambridge

University Press.

Hirshman, E. (1988). The expectation-violation effect: paradoxical effects of semantic relatedness. *Journal of Memory and Language*, *27*(1), 40–58. https://doi.org/10.1016/0749-596X(88)90047-2

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception* (pp. 309–331). New York: Holt, Reinhart &Winston.

Hock, H. S., & Egeth, H. (1970). Verbal interference with encoding in a perceptual classification task. *Journal of Experimental Psychology*, *83*(2, Pt.1), 299–303. https://doi.org/10.1037/h0028512

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, *28*(1), 113–136. https://doi.org/10.1037/0096-1523.28.1.113

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: what it is and how to measure it. In C. H. Morimoto, H. O. Istance, S. N. Spencer, J. B. Mulligan, & P. Qvarfordt (Eds.), *ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 45–52). Santa Barbara, CA: ACM Press. https://doi.org/10.1145/2168556.2168563

Hommel, B. (2002). Responding to object files: automatic integration of spatial information revealed by stimulus-response compatibility effects. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *55*(2), 567–580. https://doi.org/10.1080/02724980143000361

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, *136*(2), 189–202. https://doi.org/10.1016/j.actpsy.2010.04.011

Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: contrasting features and objecthood as indexes to memory. *Cognition*, *108*(2), 533–542. https://doi.org/10.1016/j.cognition.2008.02.011

Huey, E. B. (1908). *The psychology and pedagogy of reading*. New York, NY: Macmillan.

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, *109*(1), 49–74. https://doi.org/10.1037/0096-3445.109.1.49

Hunt, R. R., & Worthen, J. B. (Eds.). (2006). *Distinctiveness and memory*. Oxford: Oxford University Press.

Hyona, J., Radach, R., & Deubel, H. (Eds.). (2003). *The mind's eye: cognitive and applied aspects of eye movement research*. Amsterdam: North-Holland Publishing Company.

Inhoff, A. W., & Weger, U. W. (2005). Memory for word location during reading: eye movements to previously read words are spatially selective but not precise. *Memory & Cognition*, *33*(3), 447–61. https://doi.org/10.3758/BF03193062

Irvin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, *5*(1–2), 127–155. https://doi.org/10.1080/713756783

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and* *action: eye movements and the visual world* (pp. 105–133). Hove: Psychology Press.

Irwin, D. E., & Zelinsky, G. J. (2002). Eye movements and scene perception: memory for things observed. *Perception & Psychophysics*, *64*(6), 882–895. https://doi.org/10.3758/BF03196793

Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, *268*(5218), 1772–1774. https://doi.org/10.1126/science.7792605

Jahn, G., & Braatz, J. (2014). Memory indexing of sequential symptom processing in diagnostic reasoning. *Cognitive Psychology*, *68*, 59–97. https://doi.org/10.1016/j.cogpsych.2013.11.002

Jiménez, R., Cárdenas, D., González-Anera, R., Jiménez, J. R., & Vera, J. (2018). Measuring mental workload: ocular astigmatism aberration as a novel objective index. *Ergonomics*, *61*(4), 506–516. https://doi.org/10.1080/00140139.2017.1395913

Jirak, D., Menz, M. M., Buccino, G., Borghi, A. M., & Binkofski, F. (2010). Grasping language - a short story on embodiment. *Consciousness and Cognition*, *19*(3), 711–720. https://doi.org/10.1016/j.concog.2010.06.020

Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1289–1314. https://doi.org/10.1037/a0026585

Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, *30*(6), 1053–1079. https://doi.org/10.1207/s15516709cog0000_86

Johansson, R., Holsanova, J., & Holmqvist, K. (2011). The dispersion of eye movements during visual imagery is related to individual differences in spatial imagery ability. In L. Carlson, T. F. Shipley, & C. Hoelscher (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 1200–1205). Austin, TX: Cognitive Science Society. https://doi.org/10.1037/a0026585

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25*(1), 236–242. https://doi.org/10.1177/0956797613498260

Johansson, R., Oren, F., & Holmqvist, K. (2018). Gaze patterns reveal how situation models and text representations contribute to episodic text memory. *Cognition*, *175*(June), 53–68. https://doi.org/10.1016/j.cognition.2018.02.016

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, *14*(6), 540–545. https://doi.org/10.1016/j.sleep.2007.08.004

Johnson-Laird, P. (1983). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115. https://doi.org/10.1207/s15516709cog0401_4

Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of working memory in mind and brain. *Current Directions in Psychological Science*, *14*(1), 2–5. https://doi.org/10.1111/j.0963-7214.2005.00323.x

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*(3), 79–89. https://doi.org/10.1016/j.cognition.2004.06.005

Kato, C., Isoda, H., Takehara, Y., Matsuo, K., Moriya, T., & Nakai, T. (1999). Involvement of motor cortices in retrieval of kanji studied by functional MRI. *NeuroReport*, *10*(6), 1335–1339. https://doi.org/10.1097/00001756-199904260-00033

Kaufman, L., & Richards, W. (1969). Spontaneous fixation tendencies for visual forms. *Perception & Psychophysics*, *5*(2), 85–88. https://doi.org/10.3758/BF03210527

Keenan, T. R., & Benjafield, J. G. (1994). An additional measure of availability derived from the Oxford English Dictionary. *Psychonomic Bulletin & Review*, *1*(2), 255–257. https://doi.org/10.3758/BF03200777

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: spatial ability's unique role. *Psychological Science*, *24*(9), 1831–1836. https://doi.org/10.1177/0956797613478615

Kent, C., & Lamberts, K. (2006). Modeling the time course of feature perception and feature information retrieval. *Journal of Memory and Language*, *55*(4), 553–571. https://doi.org/10.1016/j.jml.2006.08.010

Kent, C., & Lamberts, K. (2008). The encoding-retrieval relationship: retrieval as mental simulation. *Trends in Cognitive Sciences*, *12*(3), 92–98. https://doi.org/10.1016/j.tics.2007.12.004

Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, *6*(12), e29221. https://doi.org/10.1371/journal.pone.0029221

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, *14*(12), 7–7. https://doi.org/10.1167/14.12.7

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kirsh, D. (1995). Intelligent use of space. *Artificial Intelligence*, (73), 31–68.

Kirsh, D. (2010). Thinking with external representations. *AI & Soc*, *25*, 441–454. https://doi.org/DOI 10.1007/s00146-010-0272-8

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1207/s15516709cog1804_1

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: behavioral measures and cortical activity. *Journal of Cognitive Neuroscience*, *15*(4), 559–73. https://doi.org/10.1162/089892903321662949

Kohlisch, O., & Schaefer, F. (1996). Physiological changes during computer tasks: responses to mental load or to motor demands? *Ergonomics*, *39*(2), 213–224. https://doi.org/10.1080/00140139608964452

Körner, A., Topolinski, S., & Strack, F. (2016). Routes to embodiment. *Frontiers in Psychology*, *6*(July), 1–10. https://doi.org/10.3389/fpsyg.2015.00940

Kosslyn, S. M. (1973). Scanning visual images: some structural implications. *Perception & Psychophysics*, *14*(1), 90–94. https://doi.org/10.3758/BF03198621

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review*, *94*(2), 148–175. https://doi.org/10.1016/B978-1-4832-1446-7.50052-2

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *NeuroImage*, *6*(4), 320–334. https://doi.org/10.1006/nimg.1997.0295

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science*, *31*(4), 549–579. https://doi.org/10.1080/15326900701399897

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., … Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. *Scientific Reports*, *8*(1), 1–7. https://doi.org/10.1038/s41598-018-23197-6

Kulakova, E., Aichhorn, M., Schurz, M., Kronbichler, M., & Perner, J. (2013). Processing counterfactual and hypothetical conditionals: an fMRI investigation. *NeuroImage*, *72*, 265–271. https://doi.org/10.1016/j.neuroimage.2013.01.060

Kumcu, A., & Thompson, R. L. (2016). Spatial interference and individual differences in looking at nothing for verbal memory. In J. C. Papafragou, A., Grodner, D., Mirman, D., & Trueswell (Ed.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2387–2392). Austin, TX: Cognitive Science Society.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/web/packages/lmerTest/index.html.

Lacey, S., & Lawson, R. (Eds.). (2013). *Multisensory imagery*. New York, NY: Springer.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychonomic Bulletin & Review*, *18*(6), 1180–1188. https://doi.org/10.3758/s13423-011-0158-x

Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: the role of gaze in mental imagery and memory. *Cognition*, *131*(2), 263–283. https://doi.org/10.1016/j.cognition.2014.01.003

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: a window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, *26*(2), 207–231. https://doi.org/10.1016/S0364-0213(01)00065-9

Lakoff, G., & Johnson, M. (1980a). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980b). The metaphorical structure of the human conceptual system. *Cognitive Science*, *4*(2), 195–208. https://doi.org/10.1016/S0364-0213(80)80017-6

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought*. New York, NY: Basic Books.

Lamberts, K., Brockdorff, N., & Heit, E. (2002). Perceptual processes in matching and recognition of complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1176–1191. https://doi.org/10.1037//0096-1523.28.5.1176

Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, *136*(6), 1045–1067. https://doi.org/10.1037/a0020970

Lauer, J. E., & Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological Science*, *27*(10), 1291–1298. https://doi.org/10.1177/0956797616655977

Le Bigot, N., Passerault, J. M., & Olive, T. (2009). Memory for words location in writing. *Psychological Research*, *73*(1), 89–97. https://doi.org/10.1007/s00426-008-0135-9

LeBoutillier, N., & Marks, D. F. (2003). Mental imagery and creativity: a meta-analytic review study. *British Journal of Psychology*, *94*(1), 29–44. https://doi.org/10.1348/000712603762842084

Leitan, N. D., Williams, B., & Murray, G. (2015). Look up for healing: embodiment of the heal concept in looking upward. *PLoS ONE*, *10*(7), e0132427. https://doi.org/10.1371/journal.pone.0132427

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*(7), 963–977. https://doi.org/https://doi.org/10.1016/0042-6989(85)90207-X

Levine, D. N., Warach, J., & Farah, M. (1985). Two visual systems in mental imagery. *Neurology*, *35*(7), 1010 LP-1010.

Levine, W. H., & Klin, C. M. (2001). Tracking of spatial information in narratives. *Memory and Cognition*, *29*(2), 327–335. https://doi.org/10.3758/BF03194927

Levinson, S. C. (1992). Primer for the field investigation of spatial description and conception. *Pragmatics*, *2*(1), 5–47. https://doi.org/https://doi.org/10.1075/prag.2.1.02lev

Levinson, S. C. (2003). *Space in language and cognition: explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Li, P., Zhang, F., Tsai, E., & Puls, B. (2013). Language history questionnaire (LHQ 2.0): a new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., … Scott, S. K. (2015). Feel the noise: relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cerebral Cortex*, *25*(11), 4638–4650. https://doi.org/10.1093/cercor/bhv134

Liu, N., & Bergen, B. (2016). When do language comprehenders mentally simulate locations? *Cognitive Linguistics*, *27*(2), 181–203. https://doi.org/10.1515/cog-2015-0123

Lloyd-Jones, T. J., & Vernon, D. (2003). Semantic interference from visual object recognition on visual imagery. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(4), 563–580. https://doi.org/10.1037/0278-7393.29.4.563

Lockhart, R. S. (2000). Methods of memory research. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 45–57). Oxford: Oxford University Press.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*, *15*(4), 838–844. https://doi.org/10.3758/PBR.15.4.838

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*(1), 96–104. https://doi.org/10.1016/j.cognition.2009.09.002

Louwerse, M. M., & Jeuniaux, P. (2012). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment: debates on meaning and cognition* (pp. 309–326). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199217274.003.0015

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: the relationship between perceptual experience and surface word form. *Behavior Research Methods*, *45*(2), 516–26. https://doi.org/10.3758/s13428-012-0267-0

MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 633–642. https://doi.org/10.1037/0278-7393.15.4.633

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, *2*(11), 547–552.

Magliano, J. P., Radvansky, G. A., & Copeland, D. E. (2012). Beyond language comprehension: situation models as a form of autobiographical memory. *Higher Level Language Processes in the Brain: Inference and Comprehension Processes*, 379–392. https://doi.org/10.4324/9780203936443

Mahr, J. B., & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences*, *41*, 1–16. https://doi.org/10.1017/S0140525X17000012

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108–114. https://doi.org/10.1016/j.tics.2004.01.003

Mäntylä, T., & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, *14*(3), 365–386. https://doi.org/10.1080/13506280500347992

Marinescu, A. C., Sharples, S., Ritchie, A. C., Sánchez López, T., McDowell, M., & Morvan, H. P. (2018). Physiological parameter response to variation of mental workload. *Human Factors*, *60*(1), 31–56. https://doi.org/10.1177/0018720817733101

Markman, K. D., Klein, W. M. P., & Suhr, J. A. (2008). *Handbook of imagination and mental simulation*. Hove: Psychology Press.

Marmolejo-Ramos, F., Elosúa, M. R., Yamada, Y., Hamm, N. F., & Noguchi, K. (2013). Appraisal of space words and allocation of emotion words in bodily space. *PLoS ONE*, *8*(12), 18–20. https://doi.org/10.1371/journal.pone.0081688

Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London. B. Biological Sciences*, *262*(841), 23–81. https://doi.org/10.1098/rstb.1971.0078

Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.

Marschark, M., & Cornoldi, C. (1991). Imagery and verbal memory. In C. Cornoldi & M. A. McDaniel (Eds.), *Imagery and cognition* (pp. 133–182). New York, NY: Springer US. https://doi.org/10.1007/978-1-4684-6407-8_5

Martarelli, C. S., Chiquet, S., Laeng, B., & Mast, F. W. (2017). Using space to represent categories: insights from gaze position. *Psychological Research*, *81*(4), 721–729. https://doi.org/10.1007/s00426-016-0781-2

Martarelli, C. S., & Mast, F. W. (2011). Preschool children's eye-movements during pictorial recall. *British Journal of Developmental Psychology*, *29*(3), 425–436. https://doi.org/10.1348/026151010X495844

Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, *77*(3), 303–309. https://doi.org/10.1007/s00426-012-0439-7

Martarelli, C. S., Mast, F. W., & Hartmann, M. (2017). Time in the eye of the beholder: Gaze position reveals spatial-temporal associations during encoding and memory retrieval of future and past. *Memory & Cognition*, *45*(1), 40–48. https://doi.org/10.3758/s13421-016-0639-2

Masson, M. E. J. (2015). Toward a deeper understanding of embodiment. *Canadian Journal of Experimental Psychology*, *69*(2), 159–164. http://doi.org/10.1037/cep0000055

Mast, F. W., & Kosslyn, S. M. (2002). Eye movements during visual mental imagery. *Trends in Cognitive Sciences*, *6*(7), 271–272. https://doi.org/10.1016/S1364-6613(02)01931-9

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–24. https://doi.org/10.3758/s13428-011-0168-7

May, R. (1985). *Logical form: its structure and derivation*. Cambridge, MA: MIT Press.

McKenzie, B., Slater, A. M., Tremellen, S., & McAlpin, S. (1993). Reaching for toys through apertures. *British Journal of Developmental Psychology*, *11*(1), 47–60. https://doi.org/10.1111/j.2044-835X.1993.tb00587.x

Meghanathan, R. N., van Leeuwen, C., & Nikolaev, A. R. (2015). Fixation duration surpasses pupil size as a measure of memory load in free viewing. *Frontiers in Human Neuroscience*, *8*(January), 1063. https://doi.org/10.3389/fnhum.2014.01063

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, *22*(4), 473–500. https://doi.org/10.1080/01690960600696916

Meteyard, L., & Vigliocco, G. (2009). Verbs in space: axis and direction of motion norms for 299 English verbs. *Behavior Research Methods*, *41*(2), 565–574. https://doi.org/10.3758/BRM.41.2.565

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, *18*(17), 732–733. https://doi.org/10.1016/j.cub.2008.07.016

Meyer, B. P., & Robinson, M. D. (2004). Why the sunny side is up. *Psychological Science*, *15*(4), 243–247. https://doi.org/10.1111/j.0956-7976.2004.00659.x

Michaelian, K. (2016a). Confabulating, misremembering, relearning: the simulation theory of memory and unsuccessful remembering. *Frontiers in Psychology*, *7*(November), 1857. https://doi.org/10.3389/fpsyg.2016.01857

Michaelian, K. (2016b). *Mental time travel*. Cambridge, MA: MIT Press.

Mill, R. D., O'Connor, A. R., & Dobbins, I. G. (2016). Pupil dilation during recognition memory: isolating unexpected recognition from judgment uncertainty. *Cognition*, *154*, 81–94. https://doi.org/10.1017/CBO9781107415324.004

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Mix, K. S., Smith, L. B., & Gasser, M. (Eds.). (2010). *The spatial foundations of language and cognition*. Oxford: Oxford University Press.

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language*, *26*(2), 165–187. https://doi.org/10.1016/0749-596X(87)90122-7

Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philosophical Transactions of the Royal Society*, *364*, 1273–1280. https://doi.org/10.1098/rstb.2008.0314

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250–9. https://doi.org/10.1016/j.jneumeth.2013.10.024

Naveh-Benjamin, M. (1987). Coding of spatial location information: an automatic process? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 595–

605. https://doi.org/http://dx.doi.org/10.1037/0278-7393.13.4.595

Naveh-Benjamin, M. (1988). Recognition memory of spatial location information: another failure to support automaticity. *Memory & Cognition*, *16*(5), 437–445. https://doi.org/10.3758/BF03214224

Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*(2), 135–183. https://doi.org/10.1016/S0364-0213(80)80015-2

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, *19*(3), 113–126. https://doi.org/10.1145/360018.360022

Niendam, T. A., Carter, C. S., & Ragland, J. D. (2010). Use of eye movement monitoring to examine item and relational memory in schizophrenia. *Biological Psychiatry*, *68*(7), 610–616. https://doi.org/10.1016/j.biopsych.2010.06.001.Use

Nittono, H., Suehiro, M., & Hori, T. (2002). Word imageability and N400 in an incidental memory paradigm. *International Journal of Psychophysiology*, *44*(3), 219–229. https://doi.org/10.1016/S0167-8760(02)00002-8

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Noton, D., & Stark, L. W. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929–942. https://doi.org/10.1016/0042-6989(71)90213-6

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *97*(20), 11120–11124. https://doi.org/10.1073/pnas.97.20.11120

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–1023. https://doi.org/10.1162/08989290051137549

O'Regan, J. K. (1992). Solving the 'real' mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *46*(3), 461–488. https://doi.org/10.1037/h0084327

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973. https://doi.org/10.1017/S0140525X01000115

Örnkloo, H., & Von Hofsten, C. (2007). Fitting objects into holes: on the development of spatial cognition skills. *Developmental Psychology*, *43*(2), 404–416. https://doi.org/10.1037/0012-1649.43.2.404

Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: on the relevance of events for perceptual simulation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(4), 579–590. https://doi.org/10.1037/xlm0000318

Otero, S. C., Weekes, B. S., & Hutton, S. B. (2011). Pupil size changes during recognition memory. *Psychophysiology*, *48*(10), 1346–1353. https://doi.org/10.1111/j.1469-8986.2011.01217.x

Otten, L. J. (2007). Fragments of a larger whole: retrieval cues constrain observed neural correlates of memory encoding. *Cerebral Cortex*, *17*(9), 2030–2038. https://doi.org/10.1093/cercor/bhl111

Paivio, A. (1990). *Mental representation: a dual coding approach*. Oxford: Oxford University Press.

Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychologysychology*, *76*(1, Pt.2), 1–25. https://doi.org/http://dx.doi.org/10.1037/h0025327

Palmiero, M., Piccardi, L., Nori, R., Palermo, L., Salvi, C., & Guariglia, C. (2016). Editorial: creativity and mental imagery. *Frontiers in Psychology*, *7*(August), 1280. https://doi.org/10.3389/fpsyg.2016.01280

Papesh, M. H. (2015). Just out of reach: on the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology: General*, *144*(6), e116–e141. https://doi.org/10.1037/xge0000125

Parker, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *4*(2), 284–293. https://doi.org/10.1037/0096-1523.4.2.284

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*(2), 97–107. https://doi.org/10.1038/nrn1603

Patson, N. D., George, G., & Warren, T. (2014). The conceptual representation of number. *Quarterly Journal of Experimental Psychology*, *67*(7), 1349–1365. https://doi.org/10.1080/17470218.2013.863372

Pearson, J., & Kosslyn, S. M. (2013). Mental imagery. *Frontiers in Psychology*, *4*(April), 198. https://doi.org/10.3389/fpsyg.2013.00198

Pecher, D., Boot, I., & Van Dantzig, S. (2011). Abstract concepts: sensory-motor grounding, metaphors, and beyond. In D. Medin (Ed.), *Psychology of learning and motivation - advances in research and theory* (1st ed., Vol. 54, pp. 217–248). https://doi.org/10.1016/B978-0-12-385527-5.00007-3

Pecher, D., van Dantzig, S., & Schifferstien, H. N. J. (2009). Concepts are not represented by conscious imagery. *Psychonomic Bulletin and Review*, *16*(5), 914–919. https://doi.org/10.3758/PBR.16.5.914

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*(2), 119–124. https://doi.org/10.1111/1467-9280.t01-1-01429

Pecher, D., & Zwaan, R. A. (2010). Introduction to grounding cognition: the role of perception and action in memory, language, and thinking. In Diane, P., & Rolf A. Z.

(Eds.), *Grounding cognition* (pp. 1–7). https://doi.org/10.1016/B978-0-240-81203-8.00002-7

Pena, M. (2002). Signal-driven computations in speech processing. *Science*, *298*(5593), 604–607. https://doi.org/10.1126/science.1072901

Perky, C. W. (1910). An experimental study of imagination. *The American Journal of Psychology*, *21*(3), 422–452. https://doi.org/10.2307/1413350

Pezdek, K., Roman, Z., & Sobolik, K. G. (1986). Spatial memory for objects and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 530–537. https://doi.org/10.1037/0278-7393.12.4.530

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2013). Computational grounded cognition: a new alliance between grounded cognition and computational modeling. *Frontiers in Psychology*, *3*(January), 612. https://doi.org/10.3389/fpsyg.2012.00612

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York, NY: Basic Books.

Pinker, S. (1995). *The language instinct: how the mind creates language*. New York, NY: Harper Perennial Modern Classics.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, *5*(April), 359. https://doi.org/http://doi.org/10.3389/fpsyg.2014.00359

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*(2), 253–279; discussion 280–336. https://doi.org/10.1017/S0140525X9900182X

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(July), 576–582. https://doi.org/10.1038/2201358a0

Purves, D., Augustine, G. J., & Fitzpatrick, D. (2001). Types of eye movements and their functions. In D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O McNamara, & S. M. Williams (Eds.), *Neuroscience* (2nd ed.). Sunderland, MA: Sinauer Associates.

Putnam, H. (1960). Minds and machines. In S. Hook (Ed.), *Dimensions of mind: a symposium* (pp. 148–179). New York, NY: New York University Press.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin*, *80*(1), 327–358. https://doi.org/10.1037/h0061470

Pylyshyn, Z. W. (1986). *Computation and cognition: toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: a sketch of the

FINST spatial-index model. *Cognition*, *32*(1), 65–97. https://doi.org/10.1016/0010-0277(89)90014-0

R Core Team, X. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Pollatsek, A. (1989). *Psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157. https://doi.org/http://dx.doi.org/10.1037/0033-295X.105.1.125

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: analyses and theoretical implications. *Cognitive Science*, *31*(1), 157–168. https://doi.org/10.1080/03640210709336988

Renkewitz, F., & Jahn, G. (2012). Memory indexing: a novel method for tracing memory processes in complex cognitive tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1622–1639. https://doi.org/10.1037/a0028073

Richardson, D. C., Altmann, G. T. M., Spivey, M. J., & Hoover, M. A. (2009). Much ado about eye movements to nothing: a response to Ferreira et al.: taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *13*(6), 235–236. https://doi.org/10.1016/j.tics.2009.02.006

Richardson, D. C., Dale, R., & Spivey, M. J. (2007). Eye movements in language and cognition: a brief introduction. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 323–344). Amsterdam: John Benjamins. https://doi.org/https://doi.org/10.1075/hcp.18.21ric

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, *133*(1), 46–62. https://doi.org/10.1037/0096-3445.133.1.46

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: looking at things that aren't there anymore. *Cognition*, *76*(3), 269–295. https://doi.org/10.1016/S0010-0277(00)00084-6

Richardson, D. C., & Spivey, M. J. (2004). Eye tracking: characteristics and methods. In G. Wnek & G. Bowlin (Eds.), *Encyclopedia of biomaterials and biomedical engineering* (Vol. 2, pp. 568–572). New York, NY: Marcel Dekker. https://doi.org/10.1081/E-EBBE-120013920

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*(5), 767–780. https://doi.org/10.1016/S0364-0213(03)00064-8

Richardson, D. C., Spivey, M. J., Edelman, S., & Naples, A. J. (2001). "Language is spatial": experimental evidence for image schemas of concrete and abstract verbs. In J. Moore & K. Stenning (Eds.), *23rd Annual Conference of the Cognitive Science Society* (pp. 845–860). Austin, TX: Cognitive Science Society.

Richardson, J. T. E. (1975). Concreteness and imageability. *Quarterly Journal of Experimental Psychology*, *27*(2), 235–249. https://doi.org/10.1080/14640747508400483

Richardson, J. T. E. (1979). Mental imagery, human memory, and the effects of closed head injury. *The British Journal of Social and Clinical Psychology*, *18*(3), 319–27. https://doi.org/https://doi.org/10.1111/j.2044-8260.1979.tb00342.x

Richmond, J. L., Zhao, J. L., & Burns, M. A. (2015). What goes where? Eye tracking reveals spatial relational memory during infancy. *Journal of Experimental Child Psychology*, *130*, 79–91. https://doi.org/10.1016/j.jecp.2014.09.013

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: evidence from eye tracking. *Developmental Science*, *12*(4), 549–556. https://doi.org/10.1111/j.1467-7687.2009.00795.x

Risko, E. F., & Dunn, T. L. (2015). Storing information in-the-world: metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, *36*, 61–74. https://doi.org/10.1016/j.concog.2015.05.014

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Ross, M., & Schryer, E. (2015). Outsourcing memory in response to an aging population. *Perspectives on Psychological Science*, *10*(6), 716–720. https://doi.org/10.1177/1745691615605258

Rubin, D. C. (1983). Associative asymmetry, availability, and retrieval. *Memory & Cognition*, *11*(1), 83–92. https://doi.org/10.3758/BF03197665

Rubin, D. C., & Friendly, M. (1986). Predicting which words get recalled: measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory & Cognition*, *14*(I), 79–94. https://doi.org/10.3758/BF03209231

Rubin, D. C., & Talarico, J. M. (2009). A comparison of dimensional models of emotion: evidence from emotions, prototypical events, autobiographical memories, and words. *Memory*, *17*(8), 802–808. https://doi.org/10.1080/09658210903130764

Rugg, M. D., Johnson, J. D., Park, H., & Uncapher, M. R. (2008). Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. In *Progress in brain research* (Vol. 169, pp. 339–352). https://doi.org/10.1016/S0079-6123(07)00021-0

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*(7), 313–319. https://doi.org/10.1016/S1364-6613(03)00131-1

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, *11*(6), 454–461. https://doi.org/10.1111/1467-9280.00288

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory (Hove, England)*, *15*(5), 508–525. https://doi.org/10.1080/09658210701391022

Salvi, C., & Bowden, E. M. (2016). Looking for creativity: where do we look when we look for new ideas? *Frontiers in Psychology*, *7*(February), 161. https://doi.org/10.3389/fpsyg.2016.00161

Santiago, J., Lupiáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512–516. https://doi.org/10.3758/BF03194099

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657–661. https://doi.org/10.1080/08995600802554748

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, *76*(4), 677–694. https://doi.org/10.1016/j.neuron.2012.11.001

Schacter, D. L., Guerin, S. A., & St. Jacques, P. L. (2011). Memory distortion: an adaptive perspective. *Trends in Cognitive Sciences*, *15*(10), 467–474. https://doi.org/10.1016/j.tics.2011.08.004

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, *15*(2), 123–148. https://doi.org/10.1007/BF00995674

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*(6), 523–542. https://doi.org/10.3758/BF03197149

Scholz, A., Klichowicz, A., & Krems, J. F. (2018). Covert shifts of attention can account for the functional role of "eye movements to nothing". *Memory and Cognition*, *46*(2), 230–243. https://doi.org/10.3758/s13421-017-0760-x

Scholz, A., Mehlhorn, K., & Krems, J. F. (2011). Looking at nothing diminishes with practice. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1070–1075). Austin, TX: Cognitive Science Society.

Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research*, *80*(1), 149–158. https://doi.org/10.1007/s00426-014-0639-4

Scholz, A., von Helversen, B., & Rieskamp, J. (2015). Eye movements reveal memory processes during similarity- and rule-based decision making. *Cognition*, *136*, 228–246. https://doi.org/10.1016/j.cognition.2014.11.019

Schönpflug, W. (1986). The trade-off between internal and external information storage. *Journal of Memory and Language*, *25*(6), 657–675. https://doi.org/10.1016/0749-596X(86)90042-2

Schubert, T. W. (2005). Your highness: vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology*, *89*(1), 1–21. https://doi.org/10.1037/0022-3514.89.1.1

Schulman, A. I. (1967). Word length and rarity in recognition memory. *Psychonomic Science*, *9*(4), 211–212. https://doi.org/10.3758/BF03330834

Schulman, A. I. (1983). Maps and memorability. In D. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 359–367). New York: Plenum Press.

Schwanenflugel, P. J. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Lawrence Erlbaum Associates.

Secora, K., & Emmorey, K. (2015). The action-sentence compatibility effect in ASL: the role of semantics vs. perception. *Language and Cognition*, *7*(2), 305–318. https://doi.org/10.1017/langcog.2014.40

Seidenberg, M. S. (2002). Does grammar start where statistics stop? *Science*, *298*(5593), 553–554. https://doi.org/10.1126/science.1078094

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321. https://doi.org/10.1016/j.tics.2008.04.008

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*(2), 297–312. https://doi.org/10.1080/01690960600732430

Seymour, P. H. K. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology*, *29*(2), 245–265. https://doi.org/10.1080/14640747708400601

Shanton, K., & Goldman, A. (2010). Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 527–538. https://doi.org/10.1002/wcs.33

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*(5), 679–687. https://doi.org/10.1111/j.1469-8986.2008.00681.x

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *The Journal of Applied Psychology*, *51*(3), 300–304. https://doi.org/10.1037/h0020586

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Slotnick, S. D. (2004). Visual memory and visual perception recruit common neural substrates. *Behavioral and Cognitive Neuroscience Reviews*, *3*(4), 207–221. https://doi.org/10.1177/1534582304274070

Smith, C. N., Hopkins, R. O., & Squire, L. R. (2006). Experience-dependent eye movements, awareness, and hippocampus-dependent memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(44), 11304–11312. https://doi.org/10.1523/JNEUROSCI.3071-06.2006

Smith, M. J., Gorman, J. M., Malaspina, D., Sharif, Z., & Amador, X. (2000). Memory and schizophrenia: differential link of processing speed and selective attention with two levels of encoding. *Journal of Psychiatric Research*, *34*(2), 121–127.

https://doi.org/https://doi.org/10.1016/S0022-3956(99)00050-3

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: a review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Sobotka, S., Nowicka, A., & Ringo, J. L. (1997). Activity linked to externally cued saccades in single units recorded from hippocampal, parahippocampal, and inferotemporal areas of macaques. *Journal of Neurophysiology*, *78*(4), 2156–2163. https://doi.org/10.1152/jn.1997.78.4.2156

Speed, L. J., & Vigliocco, G. (2014). Eye movements reveal the dynamic simulation of speed in language. *Cognitive Science*, *38*(2), 367–382. https://doi.org/10.1111/cogs.12096

Speed, L. J., & Vigliocco, G. (2016). Perceptual simulation of space, speed and time in language. In Y. Coello & M. Fischer (Eds.), *Conceptual and interactive embodiment: foundations of embodied cognition* (Vol. 2, pp. 29–45). New York, NY: Routledge. https://doi.org/10.4324/9781315751962

Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activations neural representations of visual and motor experiences. *Psychological Science*, *20*(8), 289–299.

Spering, M., & Carrasco, M. (2015). Acting without seeing: Eye movements reveal visual processing without awareness. *Trends in Neurosciences*, *38*(4), 224–258. https://doi.org/10.1016/j.tins.2015.02.002

Spivey, M. J., & Geng, J. J. (2000). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, *65*(4), 235–241. https://doi.org/10.1007/s004260100059

Spivey, M. J., Richardson, D. C., & Fitneva, S. A. (2004). Thinking outside the brain: spatial indices to visual and linguistic information. In J. Henderson & F. Ferreira (Eds.), *Interfacing language, vision, and action: eye movements and the visual world* (pp. 161–190). San Diego: Psychology Press. https://doi.org/10.4324/9780203488430

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Stark, L. W., & Ellis, S. R. (1981). Scanpaths revisited: Cognitive models direct active looking. In D. F. Fisher (Ed.), *Eye movements: cognition and visual perception* (pp. 93–226). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stocker, K., Hartmann, M., Martarelli, C. S., & Mast, F. W. (2015). Eye movements reveal mental looking through time. *Cognitive Science*, *40*(7), 1648–1670. https://doi.org/10.1111/cogs.12301

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *The Behavioral and Brain Sciences*, *30*(3), 299-313; discussion 313–51. https://doi.org/10.1017/S0140525X07001975

Szpunar, K. K. (2010). Episodic future thought: an emerging concept. *Perspectives on Psychological Science*, *5*(2), 142–162. https://doi.org/10.1177/1745691610362350

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: theory, research, and application* (pp. 225–282). New York, NY: Plenum Press. https://doi.org/10.1007/978-1-4615-9325-6

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from route and survey information. *Journal of Memory and Language*, *31*(4), 261–292. https://doi.org/10.1016/0749-596X(92)90014-O

Taylor, J. E. T., Lam, T. K., Chasteen, A. L., & Pratt, J. (2015). Bow your head in shame, or, hold your head up with pride: Semantic processing of self-esteem concepts orients attention vertically. *PLoS ONE*, *10*(9), e0137704. https://doi.org/10.1371/journal.pone.0137704

Tek, C., Gold, J., Blaxton, T., Wilk, C., McMahon, R. P., & Buchanan, R. W. (2002). Visual perceptual and working memory impairments in schizophrenia. *Archives of General Psychiatry*, *59*(2), 146–153. https://doi.org/10.1001/archpsyc.59.2.146

Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica*, *132*(2), 106–114. https://doi.org/10.1016/j.actpsy.2009.01.005

Thomas, J. G., Milner, H. R., & Hanerlandt, K. F. (2003). Forward and backward recall: different response time pattern, same retrieval order. *Psychological Science*, *14*(2), 169–174. https://doi.org/https://doi.org/10.1111/1467-9280.01437

Thomas, L. E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition*, *111*(2), 168–174. https://doi.org/10.1016/j.cognition.2009.01.005

Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, *23*(2), 207–245. https://doi.org/10.1016/S0364-0213(99)00004-X

Thomas, N. J. T. (2018a). Mental imagery. In E. N. Zalta (Ed.), *The Stanford encyclopaedia of philosophy*. Retrieved from https://plato.stanford.edu/entries/mental-imagery/index.html

Thomas, N. J. T. (2018b). The Perky experiment. In E. N. Zalta (Ed.), *The Stanford encyclopaedia of philosophy*. Retrieved from https://plato.stanford.edu/entries/mental-imagery/perky-experiment.html

Tong, F. (2013). Imagery and visual working memory: one and the same? *Trends in Cognitive Sciences*, *17*(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

Topolinski, S., & Strack, F. (2009). Motormouth: Mere exposure depends on stimulus-specific motor simulations. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(2), 423–433. https://doi.org/10.1037/a0014504

Tuan Pham, M., Meyvis, T., & Zhou, R. (2001). Beyond the obvious: chronic vividness of imagery and the use of information in decision making. *Organizational Behavior and Human Decision Processes*, *84*(2), 226–253. https://doi.org/10.1006/obhd.2000.2924

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 830–846. https://doi.org/10.1037/0096-1523.24.3.830

Tversky, B. (2009). Spatial cognition: embodied and situated. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 201-216). New York, NY: Cambridge University Press.

Tversky, B., & Lee, P. U. (1998). How space structures language. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition. Lecture notes in computer science* (Vol. 1404, pp. 157–175). Berlin: Springer. https://doi.org/10.1007/3-540-69342-4_8

Tye, M. (1991). *The imagery debate*. Cambridge, MA: MIT Press.

Ucros, C. G. (1989). Mood state-dependent memory: a meta-analysis. *Cognition and Emotion*, *3*(2), 139–169. https://doi.org/10.1080/02699938908408077

Underwood, G., & Everatt, J. (1992). The role of eye movements in reading: some limitations of the eye-mind assumption. *Advances in Psychology*, *88*(C), 111–169. https://doi.org/10.1016/S0166-4115(08)61744-6

Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: eye movements during encoding and recognition. *Visual Cognition*, *17*(6–7), 812–834. https://doi.org/10.1080/13506280902771278

Valuch, C., Becker, S. I., & Ansorge, U. (2013). Priming of fixations during recognition of natural scenes. *Journal of Vision*, *13*(3), 1–22. https://doi.org/10.1167/13.3.3.doi

van der Heijden, A. H. C. (1993). The role of position in object selection in vision. *Psychological Research*, *56*(1), 44–58. https://doi.org/10.1007/BF00572132

Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience and Biobehavioral Reviews*, *30*(5), 666–679. https://doi.org/10.1016/j.neubiorev.2005.12.001

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: a review. *Psychonomic Bulletin and Review*, (February), 1–11. https://doi.org/10.3758/s13423-018-1432-y

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience*, *32*(30), 10408–10412. https://doi.org/10.1523/JNEUROSCI.1337-12.2012

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x

Vankov, I. (2011). The interplay of spatial attention and memory: evidence for memory-driven spatial compatibility effects, but not for a facilitating role of attention in memory retrieval. *Attention, Perception, & Psychophysics*, *73*(7), 2124–2129. https://doi.org/10.3758/s13414-011-0157-z

Verges, M., & Duffy, S. (2009). Spatial representations elicit dual-coding effects in mental imagery. *Cognitive Science*, *33*(6), 1157–1172. https://doi.org/10.1111/j.1551-6709.2009.01038.x

Vermeulen, N., Niedenthal, P. M., & Luminet, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cognitive Science*, *31*(1), 183–192. https://doi.org/10.1080/03640210709336990

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews*, *35*(3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007

Vo, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: introducing the pupil old/new effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Vredeveldt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eyeclosure helps memory by reducing cognitive load and enhancing visualisation. *Memory & Cognition*, *39*(7), 1253–1263. https://doi.org/10.3758/s13421-011-0098-8

Vukovic, N., & Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychologica*, *145*(1), 98–103. https://doi.org/10.1016/j.actpsy.2013.11.002

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., … Buckner, R. L. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*(5380), 1188–1191. https://doi.org/10.1126/science.281.5380.1188

Wantz, A. L., Martarelli, C. S., Cazzoli, D., Kalla, R., Müri, R., & Mast, F. W. (2016). Disrupting frontal eye-field activity impairs memory recall. *NeuroReport*, *27*(6), 374–378. https://doi.org/10.1097/WNR.0000000000000544

Wantz, A. L., Martarelli, C. S., & Mast, F. W. (2015). When looking back to nothing goes back to nothing. *Cognitive Processing*, *17*(1), 105–114. https://doi.org/10.1007/s10339-015-0741-6

Wassenburg, S. I., & Zwaan, R. A. (2010). Readers routinely represent implied object rotation: the role of visual experience. *Quarterly Journal of Experimental Psychology*, *63*(9), 1665–1670. https://doi.org/10.1080/17470218.2010.502579

Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin and Review*, *15*(2), 426–430. https://doi.org/10.3758/PBR.15.2.426

Wenzel, M. A., Golenia, J. E., & Blankertz, B. (2016). Classification of eye fixation related potentials for variable stimulus saliency. *Frontiers in Neuroscience*, *10*(February), 23. https://doi.org/10.3389/fnins.2016.00023

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125–11129. https://doi.org/10.1073/pnas.97.20.11125

White, B. W. (1969). Interference in identifying attributes and attribute names. *Perception & Psychophysics*, *6*(3), 166–168. https://doi.org/10.3758/BF03210086

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499. http://arxiv.org/pdf/1308.5499.pdf

Wishart, J. G., & Bower, T. G. R. (1982). The development of spatial understanding in infancy. *Journal of Experimental Child Psychology*, *33*(3), 363–385. https://doi.org/10.1016/0022-0965(82)90053-4

Wühr, P., & Ansorge, U. (2007). A Simon effect in memory retrieval: evidence for the response-discrimination account. *Psychonomic Bulletin & Review*, *14*(5), 984–988. https://doi.org/10.3758/BF03194132

Xue, G. (2018). The neural representations underlying human episodic memory. *Trends in Cognitive Sciences*, *22*(6), 544–561. https://doi.org/10.1016/j.tics.2018.03.004

Yang, J., & Shu, H. (2016). Involvement of the motor system in comprehension of non-literal action language: a meta-analysis study. *Brain Topography*, *29*(1), 94–107. https://doi.org/10.1007/s10548-015-0427-5

Yarbus, A. L. (1967). Eye movements and vision. *Neuropsychologia*, *6*(4), 222. https://doi.org/10.1016/0028-3932(68)90012-2

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, *105*(1), 229–236. https://doi.org/10.1016/j.cognition.2006.09.003

Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: integration and beyond. *Journal of Personality and Social Psychology*, *77*(3), 600–619. https://doi.org/10.1037/0022-3514.77.3.600

Zanolie, K., Dantzig, S. van, Boot, I., Wijnen, J., Schubert, T. W., Giessner, S. R., & Pecher, D. (2012). Mighty metaphors: behavioral and ERP evidence that power shifts attention on a vertical dimension. *Brain and Cognition*, *78*(1), 50–58. https://doi.org/10.1016/j.bandc.2011.10.006

Zhang, J. X., & Johnson, M. K. (2004). A memory-based, Simon-like, spatial congruence effect: evidence for persisting spatial codes. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *57*(3), 419–436. https://doi.org/10.1080/02724980343000404

Zlatev, J. (2012). Spatial semantics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 318-350). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199738632.013.0013

Zusne, L., & Michels, K. M. (1964). Nonrepresentational shapes and eye movements. *Perceptual and Motor Skills*, *18*, 11–20. http://doi.org/10.2466/pms.1964.18.1.11

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1196–1207. https://doi.org/10.1037/0278-7393.22.5.1196

Zwaan, R. A. (1999). Situation models: the mental leap into imagined worlds. *Current Directions in Psychological Science*, *8*(1), 15–18. https://doi.org/10.1111/1467-8721.00004

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: six replication attempts. *PLoS ONE*, *7*(12), e51382. https://doi.org/10.1371/journal.pone.0051382

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162

Zwaan, R. A., Stanfield, R., & Yaxley, R. (2002). Language comprehensders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. A., van der Stoep, N., Guadalupe, T., & Bouwmeester, S. (2012). Language comprehension in the balance: the robustness of the action-compatibility effect (ACE). *PLoS ONE*, *7*(2), e31204. https://doi.org/10.1371/journal.pone.0031204

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*(4), 954–958. https://doi.org/10.3758/BF03196557