The Shadows Are Not All the Same:
The Nature of Object Representations Depends on the Neural Circuits Involved

Dick Dubbelde

B.S. in Psychology and Biology, May 2017, University of Wisconsin – Milwaukee

A Dissertation submitted to

The Faculty of
The Columbian College of Arts and Sciences
of the George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 31, 2022

Dissertation directed by

Sarah Shomstein
Professor of Cognitive Neuroscience

The Columbian College of Arts and Sciences of The George Washington University certifies that Dick Dubbelde has passed the Final Examination for the degree of Doctor of Philosophy as of May 11, 2022. This is the final and approved form of the dissertation.

What & When: How Objects and Uncertainty Constrain Visual Attention

Dick Dubbelde

Dissertation Research Committee:

Sarah Shomstein, Professor of Psychology, Dissertation Director

Dwight Kravitz, Associate Professor of Psychology, Committee Member

Gabriela Rosenblau, Assistant Professor of Psychology, Committee Member

Dedication

This is dedicated to Anna, because the best part of everything is sharing it with you.

Acknowledgements

Here is a list of people that (1) I love and (2) have supported me through the past five years:

(1) My dissertation committee. Please say yes.

(2) All of the GWU cognitive neuroscience department which is not included below. I chose this program because of the strong community here, and I like to think that yall chose me because you thought I could contribute to that community and the overall vibe. I hope that I did. Having this community has meant so much to me, especially during the past couple years.

(3) The graduate students and post-docs that I've got to spend time with since coming here, including Breana, AJ, Joe, Ellie, Kira, Sanika, Selene, Xiaoli, Jonathan, Pat, Michelle, Samoni, Emma, Lindsay, Prachi, Chunyue, Simon, Beck, Rebeka, Darin, Gordon, Shannon, and Wenda. The fact that I just typed out that long list of names from memory is a testament to how much you all mean to me. Thank you for helping me with everything, to include maintaining my mental health through an apocalypse, and thank you for listening (or pretending to listen) to my rambling, meandering animal facts and convoluted jokes. I hope to see you again at McNasty's every year.

(4) Dwight, for helping me to discover my love of teaching and for talking with (and sometimes just to) me for hours about everything from neuroscience to superheroes, aliens, various apocalypses, TV and movies, astrophysics, infinite

AJs, bureaucratic structure, pedagogy, buttresses, and terrestrial quadruped cognition.

(5) Sarah. It was at this point in writing this that I started to tear up. Thank you so much for seeing my potential, giving me leeway when I needed it, calling me on my bullshit, and for being one of the best friends that I have ever had. ~~You've~~ You have truly made me a much better scientist and also much better person than I ever thought I could be.

(6) Anna, the very talented local artist. I certainly would not be here writing this now if it were not for your love and support. Thank you for being my Sam on this ridiculous endeavor and being here for me whenever I couldn't recall the taste of food nor the sound of water. I love you very and always.

Abstract of Dissertation

The Shadows Are Not All the Same:
The Nature of Object Representations Depends on the Neural Circuits Involved

An object representation is the pattern of neuronal activity which is evoked by the perception of an object. The work presented in this dissertation makes the argument that object representations differ depending on the neural circuitry which is recruited in their processing, resulting in perceptual and attentional differences between objects which differ across semantic categories such as size or manipulability.

In the first paper presented, we hypothesized that the real-world size of an object is encoded into the neural representation of the object, predicting that (1) the encoding of size into objects' representation affects the perception of the objects and (2) any behavioral effects elicited by the presentation size of a stimulus are also elicited by the semantic size of objects. Experiment One used a discrimination task in which targets could appear in squares of five different sizes. Responses to targets within larger boxes were slowed compared to targets within smaller boxes, demonstrating a behavioral effect of stimulus presentation size. Experiment Two used an identical task except the squares were replaced by line images of semantic objects from two groups: small objects and large objects. It was found that the responses to targets appearing within the larger objects were slowed compared to targets appearing in the smaller objects, supporting our first prediction that perception of semantic objects is affected by the semantic size of the objects. The two experiments support our second prediction that behavioral effects elicited by the presentation size of stimuli are also elicited by semantic size, supporting

our hypothesis that the real-world size of objects is encoded into the neural representation of the objects.

In the second paper presented, we hypothesized that because the dorsal and ventral visual pathways have differing proportions of magno- and parvo-cellular input, there should be behavioral differences in perceptual tasks between manipulable and non-manipulable objects. This hypothesis was tested in adults across five experiments using a gap detection task, suited to the spatial resolution of parvocellular processing, and an object flicker discrimination task, suited to the temporal resolution of magnocellular processing. Directly predicted from the cellular composition of each pathway, a strong non-manipulable object advantage was observed in gap detection, and a small manipulable object advantage in flicker discrimination. Additionally, these effects are modulated by reducing object recognition through inversion and by suppressing magnocellular processing using red light. These results establish perceptual differences between objects dependent on semantic knowledge.

The third paper presented is a perspective piece which discusses a set of attentional phenomena that are not easily accommodated within current theories of attentional selection. We call these phenomena attentional platypuses, as they allude to an observation that within biological taxonomies the platypus does not fit into either mammal or bird categories. Similarly, attentional phenomena that do not fit neatly within current attentional models suggest that current models are in need of a revision. We list a few instances of the "attentional platypuses" and then offer a new approach, that we term dynamically weighted prioritization, stipulating that multiple factors impinge onto the attentional priority map, each with a corresponding weight. The interaction between

factors and their corresponding weights determines the current state of the priority map which subsequently constrains/guides attentional allocation. We propose that this new approach should be considered as a supplement to existing models of attention, especially those that emphasize categorical organizations.

The work presented in this dissertation argues that object representations differ dependent on their semantic content, leading to differences in how objects are both perceived and attended to. The first two papers present novel hypotheses about how the semantic categories of size and manipulability are incorporated into object representations and provide experimental evidence to support the conclusion that object representations differ across the semantic categories of size and manipulability. The third paper argues that attentional effects which derive from factors such as the semantic differences between objects challenge the current theories of attentional selection, which necessitates a new view of attention such as dynamically weighted prioritization. Taken together, what we know about an object changes our representation of the object, which alters our perception of and attention to that object.

Table of Contents

Dedication ..... **iii**

Acknowledgement ..... **iv**

Abstract ..... **vi**

Table of Contents ..... **ix**

List of Figures ..... **xi**

Chapter 1: Dissertation Introduction 1

Chapter 2: Object Perception Is Modulated By Size,

Both Actual And Semantic 8

Introduction for Dissertation 8

Introduction 17

Experiment 1 19

Experiment 2 23

Discussion 27

Chapter 3: Mugs And Plants: Object Semantic Knowledge Alters Perceptual

Processing With Behavioral Ramifications 32

Introduction for Dissertation 32

Introduction 36

Method 40

Apparatus and Stimuli 41

Task 43

Experiment 1 Results 46

Experiment 2 Results 48

Experiment 3 Results 50

Experiment 4 Results 52

Experiment 5 Results 55

Discussion 57

Chapter 4: Attention and Platypuses 62

Introduction for Dissertation 62

Introduction 68

What's With the Platypus? 70

Examples of Some Platypuses 73

What Do We Learn From the Platypuses? 80

Conclusion 82

Chapter 5: Conclusion 84

References 88

List of Figures

Figure 1 34

Figure 2 38

Figure 3 56

Figure 4 61

Figure 5 65

Figure 6 68

Figure 7 71

Figure 8 87

Figure 9 96

CHAPTER 1 - Introduction

Imagine a mug. Really picture it in your mind. What color is it? How does it feel in your mind's hand? Put that to the side next to the potted pitcher plant on your desk. Next, imagine an elephant. It's huge, isn't it? Did you know elephants like to listen to music? They sway along with the beat, just like the elephant in your head is now. These mental pictures are called representations. What you accessed in your mind was your representation of a mug and your representation of the elephant. These representations are the topic of this dissertation. What are they, how are they used, and are they all the same?

With this informal definition of a representation[1] in mind, it is worth going through the ontogeny of a new representation in order to more properly define what an object representation is. The first thing is the light bouncing from the object enters your eye and activates the rod and cones receptors at the back of your retina[2]. These cells fire[3], which activates a bipolar cell, which, after integration with the activity of neighboring cells through the horizontal cells, then activates a retinal ganglion cell (Wolfe, 2006). There are many different types of these retinal ganglion cells, each of which collects and organizes the information from the rods and cones in a different way such that the cell will fire if it detects a dot, a specific level of luminance, movement, or a number of different other things (Dhande et al., 2015). For example, two of the more notable types

[1] Your representation representation.
[2] For devastatingly interesting evolutionary reasons the photoreceptors in the vertebrate eye are in the very back of the retina, necessitating that light travels through several layers of cells before reaching the receptors. Octopuses, notably, have no such weakness.
[3] Or rather, don't, but they activate a cascade of signaling which works as most neurons do starting with the neuron which the photoreceptor synapses on to. This is called the dark current.

of retinal ganglion cell[4] are the midget and the parasol retinal ganglion cells, which are named after the size of their dendritic arbors. The midget retinal ganglion cells have small dendritic arbors, enabling them to integrate information from the rods and cones over only a relatively small area of the retina. The structure of their dendritic arbor enables the midget retinal ganglion cells to be good at detecting small differences in illumination, providing a suitable foundation for the detection of object borders (Derrington & Lennie, 1984; Leonova et al., 2003; McAnany & Alexander, 2008). In contrast to the midget retinal ganglion cells, the parasol retinal ganglion cells have large dendritic arbors, enabling them to detect stimuli across a relatively large area of the visual field, creating a perfect set-up for the detection of motion (Pokorny and Smith, 1997).

The axons of the retinal ganglion cells are what form the optic nerve, carrying the information gathered by the eye to the brain. There are two main structures in the brain to which the visual information projects: the superior colliculus and the lateral geniculate nucleus of the thalamus[5] (Rosa et al., 2000; Leopold, 2012; Kaas et al., 2022). The pathway leading to and then from the superior colliculus we will be leaving for now, but lest you fear, we will return to it in Chapter 3[6]. The visual information which projects to the lateral geniculate nucleus of the thalamus is organized according to its type (if it is encoded by midget or parasol retinal ganglion cells (or other types)), and its location in the visual field, which makes the organization of information in the lateral geniculate

[4] They're notable because they are specific to primates and we are primates.
[5] While it is an interesting story involving photo-sensitive retinal ganglion cells, we will be neglecting the suprachiasmatic nucleus because I do not want to put you to sleep.
[6] It is a story which involves dinosaurs. It's important to mention this now so that I can build suspense.

nucleus retinotopic[7] (Connolly and Van Essen, 1984). The neurons in the lateral geniculate nucleus then project to area V1 in the calcarine sulcus of the occipital lobe of the cortex, which maintains the same retinotopic organization (Connolly and Van Essen, 1984; DeYoe et al., 1996; Arcaro et al., 2009, Wang et al., 2015).

It is in V1 that we can start to see the recognizable building blocks of object representations. The neurons in V1 have very strong stimulus selectivities with each cell only firing when an oriented line is in the retinotopically corresponding part of the visual field (Hubel and Wiesel, 1959; 1962). The orientation selective neurons are then arranged in functionally defined columns such that neighboring neurons' selectivities are only slightly different and cells which are deeper or shallower in the cortex have the same selectivity. The functional columns tile the entire visual field such that every visual pixel within the visual field has retinotopically corresponding neurons which respond to all possible oriented lines (Hubel and Wiesel, 1963). Perhaps a tedious solution to the problem of vision, but clearly an effective one.

From V1, it is fairly easy to see how the visual system builds more complex object representations. In the retinal ganglion cells and in the neurons of the lateral geniculate nucleus we see selectivity to visual objects like dots (Wolfe, 2006). Many of the dot selective neurons in the lateral geniculate nucleus project to a single neuron in V1 such that if those dots are arranged in a line, then the cell in V1 fires, encoding the presence of a line in the corresponding area of the visual field. This same trick underlies the transformation of information between all of the visual areas, to varying levels of

[7] For those uninitiated to visual neuroscience, retinotopic organization can be thought of as a picture. The visual information in a picture is smaller and slightly distorted from how it is in the real world, but it maintains the same relative spatial arrangements of the objects in the scene.

abstraction. For example, several orientation-selective V1 neurons project onto a single neuron in V2[8] such that if the orientations come together into a V-shape then the neuron in V2 will fire, making it a V-shape selective neuron. Then, from V2 there are several V-shape selective neurons which project to a cell in V3 which fires if those V-shapes are arranged into a diamond shape. Transformations such as these are continued through all the visual areas, following through the temporal and parietal cortices (Wolfe, 2006).

Of course, while that explains a significant part of what the visual regions do, it is not nearly so simple. Nothing about the brain is. The retinotopic visual regions were discovered, and are usually found, using functional magnetic resonance imaging (fMRI). By presenting flashing checkerboard patterns in concentric rings (for eccentricity) and angled wedges (for polar angle) covering the visual field, one can see the retinotopic organization of the early visual regions in the evoked BOLD activity of the brain. This is true up for all of the occipital visual regions, and all of the parietal visual regions, but only for the more caudal temporal visual regions (Connolly and Van Essen, 1984; DeYoe et al., 1996; Arcaro et al., 2009, Wang et al., 2015). As you move rostrally into the visual temporal cortex the retinotopic organization becomes less clear (Kravitz et al., 2013). What you start to find is regions which are highly selective to specific sorts of visual objects such as body parts (Downing et al., 2001), common objects (Kourtzi et al., 2000), places (Epstein et al., 1998) and most notably, faces (Kanwisher et al., 1997).

In my lectures this is where I would go through the research supporting the existence of these highly specific areas, trap my students into arguing that these regions evolved for relevant categories of objects, and then stun them with the fact of the visual

[8] The second visual area, very cleverly named.

word form area which is highly selective to written letters (Cohen et al., 2000). The punchline of this pedagogical feint is that the visual areas of the temporal cortices are still retinotopic, but their selectivities are so precise as to no longer respond to flashing checkerboard patterns (Silson et al., 2015, Silson et al., 2016). Buttressing this idea is the fact that the temporal cortex is replete with 'elaborate cells' which are very similar to the orientation selective cells of V1, but their selectivities are very complex and unique[9] (Tanaka et al., 1991). An exemplar elaborate cell had selectivity to a circular shape with horizontal stripes on the top half and vertical on the bottom. As this pattern was experimentally degraded by removing parts of the pattern, tilting the circle, or changing the size, the firing rate of the cell decreased. In another interesting parallel to the orientation-selective neurons of V1, the elaborate cells are also arranged into functional columns, with other neurons deeper or shallower in the column responding to very similar stimuli or to the composing parts of the stimulus (Ito et al., 1995).

This has been a high-speed tour of the temporal visual system as it would be encountered by information first entering through the eyes, but the visual process does not end at the anterior pole of the temporal lobe. Each of these regions is not only connected to the next region in the visual pathway. Most of these regions are reciprocally connected to their preceding region, meaning an area like V2 doesn't just connect to V3, but also back to V1[10] (Kravitz et al., 2011; Kravitz et al., 2013). Keeping this fact of reciprocal connectivity in mind, it is also important to note that the process of long-term potentiation[11] is at play at every synapse in this chain of neurons (Hebb, 1949; Andersen

[9] One elaborate cell described in the original paper responded to the bottom of a potato. I hypothesize that I have many such cells, possibly only those cells.
[10] And so on for most other pairs of visual areas one can think of.
[11] Neurons that fire together wire together. When one neuron fires and causes another to fire those two neurons become more strongly associated such that it becomes easier for the first neuron to affect the second.

and Lømo, 1966). Combined with the reciprocal connections between these regions, what happens is that when visual information, first encoded cortically in V1, feeds forward through the visual areas to activate an elaborate cell you then have the elaborate cell's activity feeding back to bolster the activity of the neurons in the earlier visual regions. This circuit of neurons, through long-term potentiation, then becomes more likely to be activated upon the detection of similar visual information. This is the neural instantiation of an object representation[12], a pattern of neural activity which encodes the perception of a visual object that is reinforced and maintained by activity-driven changes in the structure of the neurons which support the pattern of activity.

Now, the topic at hand, are all object representations the same?[13] In the work presented throughout this dissertation I have found evidence that object representations differ depending on the circuits which sustain them, the size of those circuits, the histological properties of the cells which compose those circuits, and I have found evidence that the differences between object representations have repercussions for how we both perceive the corresponding objects and how we interact with them. In the first study (Dubbelde and Shomstein, in preparation), I provide evidence that the real-world size of an object becomes encoded in its neural representation due to the average size that the object is seen in retinotopic space. This study furthers our understanding of how size is perceived and how it alters how we interact with objects in our environment. The second study provides evidence that not only the basic properties of an object matter to the object's representation but the subsequent actions that we take on the object also have implications for how the object is perceived and consequently acted upon (Dubbelde and

[12] Yeehaw.
[13] No.

Shomstein, in press). This study furthers our understanding of how we relate to objects in our environment and how our goals and adaptive actions as an organism shape our perceptions and behaviors. Lastly, I present a reconceptualization of how we as cognitive neuroscientists think about attention, using the evidence provided in the studies presented here to build the case that the current frameworks for categorizing attentional effects no longer support continued development in our understanding of visual attention (Shomstein, Zhang, and Dubbelde, 2022).

CHAPTER 2 - OBJECT PERCEPTION IS MODULATED BY SIZE, BOTH ACTUAL AND SEMANTIC

One dimension in which object representations seem to differ is in size. If a visual stimulus is retinotopically larger then it encompasses a greater number of neurons throughout the visual system as it is processed. This is clearly seen in retinotopy in which the larger concentric patterns evoke activity over a larger area of the early visual cortices than the smaller concentric circles. This is significant because it has been shown that the presentation size of an object can affect how that object is perceived. To give a few examples, size seems to be intimately connected to many perceptual dimensions which are known to be processed early in their perceptual hierarchies and have dedicated neural circuitry underlying their perception. The velocity of large objects is often perceived as slower than reality while smaller objects are perceived as being faster than they really are (Brown, 1931). The size we assume an object to be also directly affects how far away we judge that object to be (Baird, 1963) and how much force to apply to lift it, even if the weight between two differently sized objects is equal (Buckingham & Goodale, 2010). Beyond vision, auditory perceptions also seem to have an implicit connection to size. Takeshima and Gyoba (2013) have shown that participants will judge visual stimuli to be larger than they appear when they are paired with louder auditory stimuli than with quieter (Takeshima & Gyoba, 2013). These studies evidence the idea that the size of stimuli is intrinsic to our perception of them inasmuch as size affects other very basic properties of object perception.

The effect of size not only applies to the presentation size of stimuli, but also to the known size of a semantic object (semantic size)[14]. When presented with playing cards of different sizes, participants reported small playing cards as being further away and the larger playing cards as being closer (Ittelson, 1951). Without postulating the existence of a semantic size for a playing card, it's difficult to imagine why participants would make this error. A similar effect is seen in studies of mental imagery. When asked to imagine animals of different sizes, larger animals are imagined at shallower visual angles, meaning they are imagined further away than the smaller animals, presumably to accommodate the extra size (Kosslyn et al., 1978; Hubbard & Baird, 1988). In a particularly illustrative study, two-year-olds were shown to make size errors (DeLoache, 2004). The children were allowed to play in a room with several large toys including a slide, a chair, and one of those classic red and yellow cars. Then they would be presented with miniature versions of those toys. The children who were between 20 and 24 months would sometimes make size errors in which they would try to slide down the miniature slide or get into the miniature car the same way that they would with the larger objects[15]. These children were accessing their representations of these objects, and their memory of these objects, and making the assumption that the miniature toys would behave the same way as the larger equivalents. Interestingly, the size errors would taper off after 24 months, suggesting that more mature neural circuits are more able to integrate the children's representations of the objects with their current perception. These results taken

[14] For example, remember our musical elephant. We know that elephants are large, when we see them they evoke activity across a large area in our visual cortices, and so elephants are semantically large, regardless of their musicality.

[15] Interestingly, the childrens' grip apertures' for opening the car door were calibrated to the actual size of the small car and not the semantic size of the car. This foreshadows a concept which is introduced in the next section, the two visual streams.

together, it seems that size is an implicit part of how we represent and thus perceive semantic objects.

Semantic size has been quantified by asking participants to draw various objects (Konkle & Oliva, 2011). It was found that larger objects were consistently drawn larger on the paper. In two follow-up studies, one in which participants imagined an object on a computer screen and then resized a bounding box, and one in which participants resized actual object images, a similar effect was found. Large objects were imagined larger and they were resized to be larger. A later study (Long et al., 2016) followed up on these findings by using a visual search paradigm. Objects were split into two groups, large and small, and participants were to find a given object amongst a field of either similarly sized or differently sized objects. It was found that each object was easier to find (faster RTs) when it was among differently sized objects. When objects were among a field of like-sized objects they no longer differed in a searchable dimension and thus search fell from parallel to serial. Knowing that all of the objects were equally sized retinotopically, this means that semantic size is a searchable dimension, implying that it is operationalized at a fairly low level in the visual hierarchy. Interestingly, two control experiments were run. First, objects were transformed to 'texforms' which preserve the low- and mid-level features of the objects but renders them unidentifiable. The same pattern of results, albeit diminished, was found in which texforms derived from small objects were more quickly found when among a field of texforms derived from large objects and vice versa. Secondly, when the objects were fully scrambled such that the mid-level features were no longer preserved, object group no longer had any bearing on the search slopes. Together, these results show that size is an intrinsic aspect of our

representations of objects and that it is bound not only to the semantic knowledge of an object, but also to the mid-level features which compose the object.

Semantic size can change how far we perceive an object to be, where in our environment we imagine images, how quickly we can search through visual scenes, and it can affect how quickly we are able to attend to and shift across objects, but whence this semantic knowledge of size? Semantic size is an organizing factor for visual representation throughout the temporal cortex (Konkle & Caramazza, 2013). Smaller objects evoke BOLD more in the ventral areas of the temporal cortex whereas the larger objects evoked more dorsal regions. Also, the texform versions of object images yield the same pattern of BOLD activation, although again to a lesser degree than the full images (Long et al., 2018). Importantly, these broad areas are consistent with the well-known Kanwisherian areas (Kanwisher et al., 1997; Epstein & Kanwisher, 1998), with FFA being in the animal zone and PPA being in the large object zone. From these studies it can be seen that size is an organizing feature, predicting where across the temporal cortex an object will evoke the largest change in BOLD activity, and it is also an intrinsic feature to the representations of visual objects, being both evoked by presenting a semantic object or in presenting a semantically obscured image which maintains the mid-level features of the object.

Delving into the cell physiology of the temporal cortex we find a suitable substrate for size calculations in the elaborate cells (Tanaka et al., 1991). One salient feature of these elaborate cells aside from their high degree of specificity is that the elaborate cells also had preferred orientations. Many of the cells would curtail their firing rates as their preferred stimulus was rotated in either direction. It's a known feature of

semantic objects, including most prominently faces (Diamond & Carey, 1986), that they are more difficult to recognize when presented in non-canonical orientations. More relevantly, the same was true in many of the elaborate cells for size (Ito et al., 1995). As their preferred stimulus was made larger or smaller the cells would curtail their firing, which is size selectivity. The size selectivity of the elaborate cells is also evidenced behaviorally. Objects are more difficult to match in memory tasks when they are not presented at the same sizes (Besner, 1983; Bundesen & Larsen, 1975; Jolicoeur, 1987). The size selectivity of the elaborate cells also explains the viability of the texforms in the visual search study of Long et al., (Long et al., 2016) and the fMRI study of visual temporal cortex (Long et al., 2018). In those studies, in which semantic size was evidenced as a searchable dimension and size was shown as an organizing feature of the visual temporal cortex, similar but diminished results were found for texforms which obscured the identity of the objects but preserved the mid-level features. The use of semantic objects and corresponding texforms in behavioral and fMRI studies together with the properties of the object selective cells in the temporal cortex demonstrate the existence of semantic size and provide a mechanism by which it is instantiated in neural circuits.

The following paper presents data supporting the hypothesis that the semantic size of an object affects the perception of the object in the same way that actual size does. The experiments in the paper demonstrate a behavioral effect which can be evoked by both presentation size and semantic size, suggesting that the two types of size may share a common mechanism. The paper however does not go into the specific about how that mechanism might work.

There are perceptually relevant representations of object size already in the early visual cortex, as evidenced by the retinotopic representation of objects in the Ponzo illusion[16] tracking the perceived size of the objects rather than the presentation size. The Ponzo illusion was presented to participants in a 3D render in which gray spheres were presented as near or far to the viewer in a brick hallway (Murray et al., 2006). Participants viewed the spheres and manipulated the size of a gray circle until it was equal to the size they perceived the sphere to be. Participants adjusted this circle to be smaller when the sphere was near in the hallway and they adjusted it to be larger when it was far in the hallway, which is the standard Ponzo result. Taking this into a scanner, the researchers first defined retinotopic concentric ROI emanating from where in the display the spheres of the Ponzo illusion would be presented. The participants were then presented with the Ponzo illusion and the BOLD response was measured for both the near and the far sphere. A strict retinotopic understanding of V1 would predict that the retinotopic area which responds to the two spheres should be equal because their display size is identical. The researchers found that the near sphere more strongly activated the smaller concentric ROIs while the far spheres more strongly activated the larger concentric ROIs. When the hallway display was removed (removing the Ponzo effect) and the actual size of the stimuli was changed, a very similar result was found. These surprising results have been replicated several times and in a number of contexts (Murray et al., 2006, Fang et al., 2008). Supporting this finding, retinal afterimages change size

[16] This is the illusion with two lines approaching a vanishing point on the horizon. Objects placed between the lines closer to the vanishing point appear to be larger than objects placed further from the vanishing point.

dependent on the surface they are projected on, a phenomenon known as Emmert's law[17] (Emmert, 1881; Sperandio et al., 2012). These results show that the V1 representation of size does not correspond to the actual size of stimuli but rather to the experienced size. This also suggests a mechanism by which size can affect the perception of an object. If size alters the amount of retinotopic space in the early visual cortices which encodes a given percept, then the perceived size of that object will also change.

How could this representational morphing be achieved in the neural circuits? Ni et al. (Ni et al., 2014) suggest that the receptive fields of the neurons shift to incorporate the interpreted size rather than the actual size. For example, for the far sphere in the Ponzo illusion, the neurons which are activated by the actual size recruit the neighboring neurons which usually would respond to an actual larger circle such that they now respond to the circle which is interpreted as being larger. Conversely, for the near sphere the opposite occurs, with the retinotopically activated neurons recruiting neurons which would normally respond to smaller circles. This would result in an object, which is interpreted to be a different size than presented, eliciting a retinotopic response in accordance with its interpreted size and not its actual size. Using the Ponzo illusion to induce perceptual size changes in macaques, they found with single cell recording that the receptive fields of V1 retinotopic neurons would indeed shift either toward the center of the object if the object were perceived larger than the actual size or the receptive fields would shift outward if the object were perceived smaller than the actual size (Ni et al., 2014). To accomplish this there must be some method by which these retinotopic neurons

[17] Try it. Stare at a bright image on your computer screen for the better part of a minute. Then look at a white wall across the room from you. The afterimage will appear larger on the wall than it does on your computer screen.

in V1 can communicate with one another. Such a thing exists in the horizontal cells (Gilbert & Wiesel, 1989). Evidence for the implication of horizontal cells in representational shifting across the early visual cortex retinotopy is garnered from a study using both the Ebbinghaus and the Ponzo illusions (Schwarzkopf et al., 2011). They found that the strength of the illusion was inversely correlated with the size of V1 as measured by a standard retinotopy paradigm in fMRI. The reason for this is that in a larger V1 more horizontal cells would be needed to cause the same representational shift, as the increased size of the V1 retinotopy corresponds to the same size of visual field leading to a greater neural density of any given visual pixel. If more horizontal cells are required because a greater cortical distance has to be covered to create the same shift then the intensity of the size illusion will be diminished.

Now, in V1 we have a suitable neural substrate for the perception of size. The retinotopic size of a stimulus in V1 tracks the perceived size of the stimulus and there is a biologically plausible mechanism for this occurring through the horizontal cells. We also have semantic size being stored in the temporal visual cortices at the neuronal level with the elaborate cells' size selectivity and at the neuroanatomical level with size being an organizing factor of the temporal visual cortices. Lastly, there are extensive feedback connections going from the temporal visual cortices back to the early visual areas. Taking this all together, we have a potential neural mechanism for how semantic size can affect the neural representation of a semantic object in early visual cortices such both the perception of and the attention on the semantic object is affected.

Taken together with the following paper, there is reason to believe that the semantic representations of different objects are different according to their size which

has an effect on how semantic objects are not only attended to, but also on how semantic objects are perceived.

Object Perception Is Modulated By Size, Both Actual And Semantic

Dubbelde, D., and Shomstein, S. (in preparation)

Introduction

Imagine an elephant. Elephants are quite friendly and they like to listen to music. Now, imagine a mouse. Mice are less musically inclined. You likely imagined the elephant as being much larger than the mouse. Kosslyn et al. (1978) asked their participants to imagine animals, such as elephants and mice, before asking for a report on how far away the animals were imagined. The smaller animals, like the mouse, were reported as being much closer to the participants than the large animals, like the friendly elephants, suggesting that the representations of animals have semantic size, a canonical size at which they are recalled. Studies since have shown that semantically large objects are typically drawn larger than semantically smaller objects (Konkle and Oliva, 2011), that semantic size can be used as a search dimension in classic visual search tasks (Long et al., 2016), and that objects which are presented in scenes at a size violating their semantic size are more often missed in realistic visual search tasks (Eckstein et al., 2017). Taken together, it is clear that semantic size is an important part of an object's representation.

More recently it has been shown that knowledge of semantic size directly impacts how attention interacts with object selection. It was demonstrated that an object's semantic size affects attentional shifting independent of the presentation size of the object (Collegio et al., 2019). The authors presented targets that appeared on top of a rectangular object. Following a spatial cue that captured attention the target was then presented either

at the cued end of the object or the uncued end. Objects consisted of line drawings of commonly known objects which were either semantically small (e.g., domino, credit card) or semantically large (e.g., vending machine, pool table). It was found that when the target appeared at the uncued end of the object participants were slower to respond if the object was semantically large than if it was semantically small. These results provided strong support for the hypothesis that semantic size directly impacts how attention interacts with object representations.

It has been shown that semantic size can influence the operation of attention in various tasks (e.g. visual search, spatial cuing tasks). The next question to ask is how does size become intertwined with an object's representation such that it can affect attentional operating? We hypothesize that effects elicited by low-level features, such as retinotopic size, become inextricably intertwined with semantic representations such that they become an intrinsic part of an object's representation such that they affect perception within that object. One such low-level effect elicited by retinal size is that targets in larger objects are processed less efficiently than targets in smaller objects (Castiello and Umiltà, 1990). In Castiello and Umiltà's paradigm (1990) , a square appeared to either side of a central fixation cross. After a short duration, a detection target could appear in the square. Critically, the square could appear at three different sizes, keeping the location of the target the same at each size. Targets which appeared in the larger squares were detected more slowly than targets appearing in the smaller squares with the medium sized squares yielding intermediate RTs. If the effect of stimulus size, as demonstrated by Castiello and Umiltà (1990), does become a part of an object's

representation then the same effect should be elicited by objects with equal presentation size but differing semantic size.

In the current study we test our prediction that an effect elicited by presentation size can also be elicited by semantic size in two experiments. First, we replicate Castiello and Umiltà's (1990) finding that targets are less efficiently processed in larger squares than in smaller squares. Secondly, we employ the same paradigm but replace the different sized squares with semantic objects that have equal presentation sizes but different semantic sizes. The semantic sizes were divided into two groups, large and small. Following from our hypothesis that low-level effects become intrinsic to semantic representations, we predict that targets in semantically large objects than will be less efficiently processed than targets in semantically small objects.

Experiment 1

Participants

Participants were recruited remotely from Amazon Mechanical Turk with compensation of US $7 per hour. Participants were given informed consent and reported normal or corrected-to-normal visual acuity. All experimental procedures were approved by the institutional review board of The George Washington University. The locations of the Turk workers were limited to the United States and their lifetime approval rates on the Mechanical Turk site were restricted to higher than 95%.

For Experiment 1, 72 participants were recruited. Sixty-three participants' data were analyzed after removing 2 participants for having less than 70% of trials remaining after removing all RTs below 250 ms (5.19% of trials) and above 1500 ms (0.03%),

removing 5 participants for having less than 70% accuracy, and removing 2 participants for having a mean RT more than 2 standard deviations away from the overall mean.

Stimuli and task

In Experiment 1, each trial began with a single central fixation point, which subtended a visual angle of 1° x 1° (Fig. 1A). The central fixation point was rendered in black for 500 ms on a gray background (HSV = 0. 0. 50). After 1000 the fixation presentation, a square drawn in a black outline appeared, centered at 4° to the left or right of the fixation point. The side of presentation was counterbalanced such that each participant saw an equal number of stimuli on the left and on the right. The stimuli were displayed at five different sizes (0.72º, 1.08º, 1.44º, 1.8º, and 2.16º). After 500 ms of object presentation, a Gabor patch appeared in the center of the object such that the Gabor was always in the same location regardless of the object's size. The Gabor was slanted 45º to the left or right. Participants were to indicate the direction of slant by pressing the 'm' key if the Gabor slanted to the right and by pressing the 'z' key if the Gabor slanted to the left. Participants were instructed to respond as quickly and as accurately as possible. After 300 ms of Gabor and object presentation the Gabor and the object offset. The experiment only continued to the next trial after the participant responded. If the participant was correct the experiment immediately progressed to the next trial. If the participant was incorrect the fixation turned red for 500 ms to provide negative feedback before the next trial began. Each participant did 300 trials.

Results

Experiment 1 was designed to examine the effect of object presentation size on target processing efficiency. Based on the results found by Castiello and Umiltà (1990) we predicted an effect in reaction time such that targets are less efficiently processed when in larger objects than when in smaller. Both reaction time and accuracy were analyzed.

A repeated-measures analysis of variance (ANOVA) was conducted on accuracy with object size (small, medium small, medium, medium large, and large) as the within-subject variable. Mauschly's test of sphericity indicated that the assumption of sphericity was violated ($p < 0.05$) so the Greenhouse-Geisser sphericity correct was applied. The ANOVA revealed no significant main effect of size ($F(1, 3.594) = 1.235$, $p = 0.298$, $\eta_p^2 = 0.020$) with means ($M_{small} = 95.2$, 95% CI [±2.6]; $M_{mediumsmall} = 94.8$, 95% CI [±3.2]; $M_{medium} = 95.5$, 95% CI [±2.7]; $M_{mediumlarge} = 95.4$, 95% CI [±3.0]; $M_{large} = 95.7$, 95% CI [±3.3]). Post hoc tests between each pair of conditions revealed no significant differences.

A repeated-measures analysis of variance was conducted on RT with object size (small, medium small, medium, medium large, and large) as the within-subject variable. Mauschly's test of sphericity indicated that the assumption of sphericity was violated ($p < 0.05$) so the Greenhouse-Geisser sphericity correct was applied. The ANOVA revealed a significant main effect of size ($F(1, 3.431) = 4.128$, $p = 0.005$, $\eta_p^2 = 0.062$) with means ($M_{small} = 502.441$, 95% CI [±33.623]; $M_{mediumsmall} = 498.637$, 95% CI [±33.483]; $M_{medium} = 507.337$, 95% CI [±35.417]; $M_{mediumlarge} = 506.873$, 95% CI [±33.164]; $M_{large} = 509.990$, 95% CI [±33.711]). Post hoc tests between each pair of conditions revealed a significant difference between the medium small condition and the large condition (t = -

3.627, p = 0.003). No other comparisons survived a Bonferroni correction for multiple comparisons on a family of 10.

The results of Experiment 1, albeit with a slightly different paradigm, corroborate Castiello and Umiltà's (1990) observations that targets are less efficiently processed inlarger objects than in smaller objects while no effect is found in accuracy, behaviorally demonstrating a perceptual effect elicited from the low-level feature of presentation size.

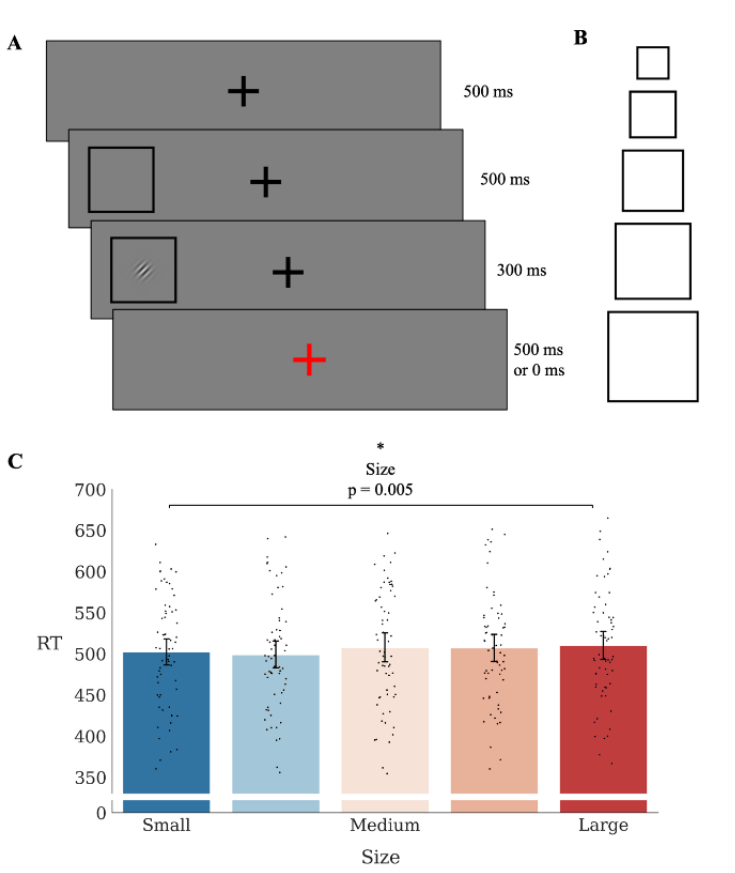


Figure 1. (A) The paradigm used in Experiment 1 with the time for each frame presented to the right. (B) The relative sizes used in Experiment 1. (C) The results from Experiment 1. The error bars depict the standard error of the mean and each point represents a participant's mean for the size condition.

Experiment 2

Participants

Participants were recruited remotely from Amazon Mechanical Turk with compensation of US $7 per hour. Participants were given informed consent and reported normal or corrected-to-normal visual acuity. All experimental procedures were approved by the institutional review board of The George Washington University. The locations of the Turk workers were limited to the United States and their lifetime approval rates on the Mechanical Turk site were restricted to higher than 95%.

For Experiment 2, 64 participants were recruited. Forty-nine participants' data were analyzed after removing 9 participants for having less than 70% of trials remaining after removing all RTs below 250 ms (7.96% of trials) and above 1500 ms (3.34%), removing 4 participants for having less than 70% accuracy, and removing 2 participants for having a mean RT more than 2 standard deviations away from the overall mean.

Stimuli and task

The design of Experiment 2 was identical to Experiment 1 except line drawings of common objects were presented instead of squares. The stimuli in Experiments 2 consisted of line drawings of real-world, everyday objects drawn both from The Noun Project, an online repository of object icons and clip art (https://thenounproject.com). During the experiment, the stimuli were displayed as large as possible in a 2.16° x 2.16° area (equivalent to the largest square in Experiment 1) and all stimuli were presented in black (HSV = 0, 0, 0) on a gray background (HSV = 0, 0, 50). All objects are displayed in Figure 2C. The object stimuli consisted of 9 line-drawings of large objects, and 9 line-

drawings of small objects. The small objects were as follows: can, cell phone, lighter, salt shaker, domino, playing card, spray paint, toothpaste, and power outlet. The large objects were: mirror, pool table, vending machine, door, throne, grandfather clock, portable toilet, kayak, and phone booth. The items in each group were adopted from Collegio et al. (2019) and the semantic size of each object was verified in a post-experiment survey completed by the same participants. In each trial of the survey, a single line drawing was presented centrally, with a six-point rating scale presented underneath. Participants were instructed to rate the size of each of the objects on a scale of one to six, one being very small and six being very large. The size ratings were used as an independent measure of each object's real-world size and participants were agnostic to the size classifications of the objects made prior to the experiments. There was a significant difference between the small and large objects in participants' size ratings in an independent samples t-test ($t(16) = 13.139$, $p < 0.001$), confirming that the participants were aware of the semantic size of each object.

Additionally, the objects were controlled for spatial frequency (average distance from origin of the points calculated from a 2D fast Fourier transform) for each object. Assessed with an independent samples t-test, there was no mean difference between object groups in spatial frequency (t-test ($t(16) = 1.811$, $p = 0.089$), suggesting that there were no low-level featural differences in the objects between the groups.

Results

Experiment 2 was designed to examine if the presentation size effect found in Experiment 1 could also be elicited by semantic size. We predicted that targets would be

less efficiently processed when in the semantically larger objects than when in the semantically smaller objects, mirroring the effect found in Experiment 1, despite identical physical size. Both reaction time and accuracy were analyzed at the group level (small or large).

A paired samples t-test was conducted on accuracy across object size (small, large), revealing no significant effect of size ($t(1, 48) = 0.007$, $p = 0.934$, $\eta_p^2 < 0.001$) with means ($M_{small} = 94.9$, 95% CI [±2.9]; $M_{large} = 94.9$, 95% CI [±2.8]). A paired samples t-test was also conducted on RT across object size (small, large), revealing a significant effect of size ($t(1, 48) = 4.899$, $p = 0.032$, $\eta_p^2 = 0.093$) with means ($M_{small} = 554.248$, 95% CI [±41.116]; $M_{large} = 558.572$, 95% CI [±40.837]).

In order to account for each participant having different levels of experience with each object, and the consequent inter-object variance in each size group, a linear mixed effects model was constructed in R using the lmerTest package (Kuznetsova et al., 2017). The statistical values were calculated using the Satterthwaite approximation. The model contained a dependent variable of RT for each trial, a fixed effect of Size, and a random effect of object nested within participant, which included both a random intercept (accounting for RT differences between participants) and a random slope (accounting for participants having different experience with each unique object). [RT~ Size + (Object | Participant)]. The LME showed a significant effect of size, with large objects having a slower average reaction time than small objects ($t(90.088) = 2.157$, $p = 0.0337$). The group level results support our hypothesis that targets are less efficiently processed when in large objects than when in semantically smaller objects, supporting our hypothesis that

behavioral effects elicited by a low-level feature, such as presentation size, can also be elicited by a semantic object which often occurs with that low-level feature.

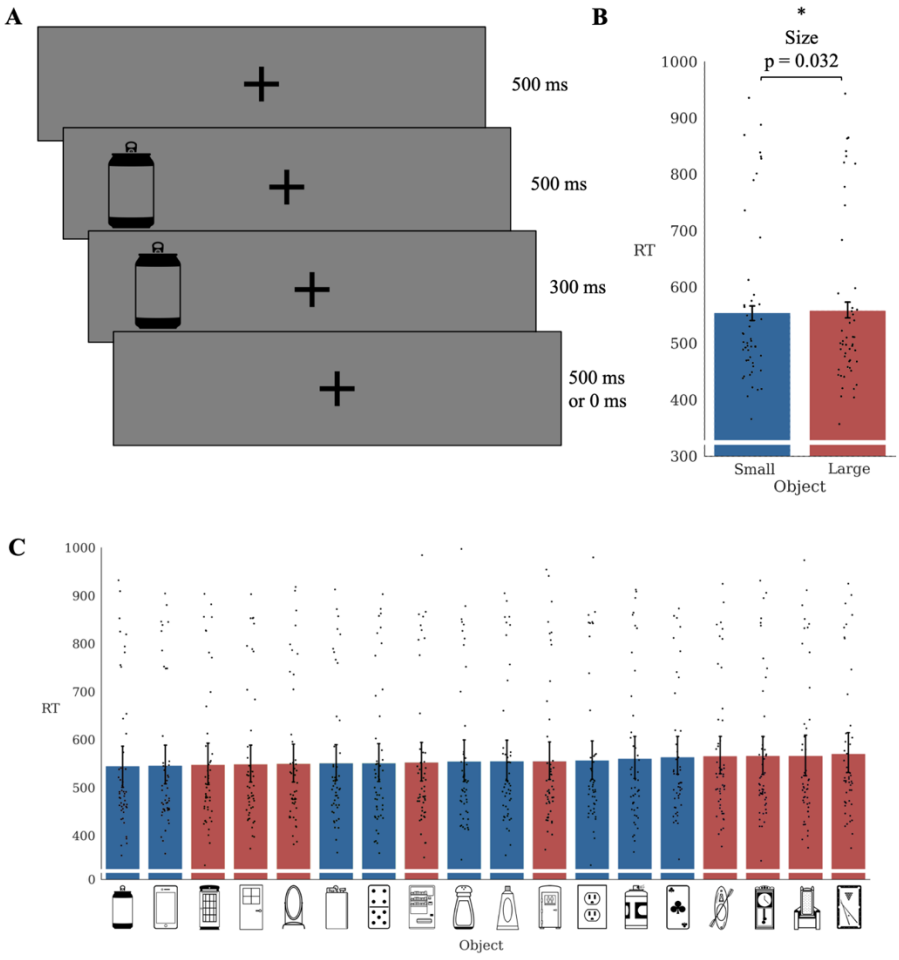


Figure 2. (A) The paradigm used in Experiment 2. (B) The results from Experiment 2 analyzed by group. The error bars depict the standard error of the mean and each point represents a participant's mean for the size condition. (C) The results from Experiment 2 depicting the data for each individual object. As in Figure 2B, semantically large objects are colored red and semantically small objects are colored blue.

Discussion

We predicted that the size effect elicited by actual retinal size would also be elicited by semantic size. Experiment 1 replicates the size effect originally shown by Castiello and Umiltà (1990) across five different sizes. Experiment 2 demonstrates that the size effect found in Experiment 1 is also elicited by semantic size at the group level with targets being less efficiently processed when in the large object group than when in the small object group. These results support our hypothesis that attentional effects elicited from low-level features, such as size, become intrinsic to semantic objects such that the behavioral effect of size can be elicited both by presentation size and semantic size.

Our findings raise an important question about the mechanism for the size effect and how it becomes intrinsic to the object representation. First, and contrary to the traditional notion of object position invariance, there is a growing amount of evidence that object representations never become independent of the retinal positions in which the objects were perceived (Kravitz et al., 2008; Kravitz et al., 2010). Presenting objects in the upper and lower halves of both hemifields, both behavioral and neurophysiological evidence was found showing that objects were primed if they occurred twice in the same quarter of the visual field, but not if they appeared twice in different quarters of the visual field (Kravitz et al., 2010). If the objects had a position invariant representation at some point in the visual hierarchy then priming should have been found regardless of object location, but the priming effect was found to be location specific, evidencing the idea that even high-level semantic representations are still grounded in location-sensitive lower-level circuits. Additionally, it has been found that the effect of location specificity

persists through time such that facial features and body parts are more speedily recognized if they appear in the visual field where they typically do (de Haas et al., 2016; Chan et al., 2010), demonstrating that the location specificity of perceived objects is not transient and it becomes associated to frequently cooccurring object representations.

The current study demonstrates that it is not just location specificity that becomes intertwined with object representations. Not only are object representations grounded to the locations in which they were perceived but if the low-level features of the object originally elicited some effect, such as the size effect of Castiello and Umilta (1990), then the low-level effect also becomes intrinsic to the object representation. A similar finding has also been found for known object manipulability in a study which demonstrated that near-hand effects (namely, higher spatial resolution for stimuli distal from the hands, higher temporal resolution for stimuli proximal to the hands) can be elicited by objects with or without semantic manipulability (Dubbelde and Shomstein, 2022). In the same way that the effect of presentation size in the current study becomes associated to object representations, the effect of hand placement also becomes associated to object representations such that it can be elicited with semantic objects regardless of hand placement.

The hypothesis that low-level effects become intrinsic to semantic object representations can be further tested by investigating if other lower-level effects extend to frequently cooccurring semantic objects. For example, it has been demonstrated that attentional shifting through objects (such as in the Egly paradigm (Egly et al., 1994)) is modulated by the orientation of the objects Barnas et al. (2016; 2019). If the effect of

orientation is also elicited from semantic objects which have canonical orientations, such as cars being typically horizontal, then the hypothesis would be further bolstered.

A future direction is ascertaining how the semantic size effect is achieved by the underlying neural circuits. We propose two distinct possibilities. The first possible mechanism is suggested by the presence of retinotopic biases throughout the ventral temporal cortices as evidenced by priming effects when objects reappear in the same location but not in a different location (Kravitz et al., 2008; 2010) and facial features and body parts being more speedily processed in the areas of the visual field in which they are most often perceived (de Haas et al., 2016, Chan et al., 2010). The idea that a similar mechanism could be at play for semantic size is buttressed by the findings that size is a macro-level organizing principle of the temporal cortex as evidenced by fMRI studies (Konkle and Caramazza, 2013) and that there is selectivity for size throughout many of the highly selective elaborate cells throughout the temporal cortex (Tanaka et al., 1991; Ito et al., 1995). With object size being coded into the object selective areas of the temporal cortex, and with the object selective areas being functionally connected to the retinotopic earlier visual areas, there is a possible neural circuit for how semantic size could become stored in object representations, activated by new perceptual input, and then fed back into the earlier visual areas to affect lower-level visual processing.

Alternatively, the second possible mechanism is that size and semantic size are relevant to the allocation of attention. It has been demonstrated that semantic information can bias attentional allocation even when it is task irrelevant such that attention is biased to semantically related objects (Malcolm et al., 2015, Shomstein et al., 2019). Supporting the significance of task-irrelevant semantic information, it has been found that

semantically related items evoked a higher change in BOLD activation relative to semantically-unrelated items in both the early visual cortices and the intraparietal sulcus, implying that the effect of semantic relatedness has an attentional origin rather than being derived from object recognition processes, as the same difference in evoked BOLD was not found in the lateral occipital complex (Nah et al., 2021). In the current study, the size of the objects, both actual and semantic size, was task irrelevant. In both experiments the target always appeared in the same absolute location at the same time interval from the onset of the fixation cross, rendering the objects, both the squares and the semantic objects, irrelevant to participants' ability to do the task. If the mechanism underlying the effect of semantic size is the attentional mechanism underlying semantic relatedness then the effect of object size should be observed at a time interval relative to the attentional selection of the objects rather than relative to the onset of the objects. By adapting the paradigm used in Experiment 2 by presenting two objects and then indicating the location of the target with a cue, object onset could be separated from object selection. Then, by jittering the time between the object onset and the cue onset, and by presenting the target at different times, the temporal window in which the semantic size effect occurs could be measured in relation to both the object onset and object selection. The temporal circuit mechanism hypothesis would be supported by the semantic size effect occurring at a consistent interval from the object onset while the attention mechanism hypothesis would be supported by the semantic size effect occurring at a consistent interval from object selection.

Previous literature has demonstrated that both presentation size and semantic size are important factors in our perception of and interaction with objects. The experiments

presented here provide evidence that both presentation size and semantic size affect the efficiency of target processing in the same way, supporting the hypothesis that semantic size effects are presentation size effects which have become encoded into objects' semantic representations. Future work investigating the nature of the semanticity of objects should continue to investigate how behavioral effects elicited by low-level features commonly associated with the object can come to be elicited by the semantic identity of the object.

CHAPTER 3 – MUGS AND PLANTS: OBJECT SEMANTIC KNOWLEDGE ALTERS PERCEPTUAL PROCESSING WITH BEHAVIORAL RAMIFICATIONS

Throughout the introduction and the first section of this dissertation we have toured the temporal vision system. There is another. The idea of two visual streams began to come into focus with Ungerleider and Mishkin's work on macaques (Ungerleider and Mishkin, 1982), noticing that macaques with lesions in the inferotemporal cortex had difficulty in discriminating between patterns whereas macaques with posterior parietal lesions had difficulty in localizing objects according to landmarks[18]. Patient DF showed the same dissociation in humans. DF was able to put an envelope into a mail slot easily, with an appropriate grip aperture and arm movement, but she could not mime the same action in response to the presentation of an oriented line. DF had damage to the temporal cortex as a result of carbon monoxide poisoning[19], which damaged her temporal visual system, leaving her parietal visual system mostly intact and functioning (Goodale et al., 1991; Goodale and Milner, 1992). Conversely, patient RV was able to easily discriminate between visual objects but had immense difficulty in manually manipulating them resulting from extensive parietal damage while having an intact temporal visual stream (Goodale et al., 1994). We now know the visual streams as the ventral "what" stream,

[18] These became known as the "what" pathway and the "where" pathway, with the temporal lesion inhibiting the macaques' ability to identify what something is and the posterior lesion inhibiting the macaques' ability to identify where something is.

[19] If you ever see a ghost you should open a window. It is likely that you are being poisoned by carbon monoxide.

which is the temporal visual system discussed in the introduction, and the dorsal "where/how" stream, the parietal visual system.

Patients in the neuropsychology literature brought the two visual streams hypothesis into focus, but it is built on decades of research into the neuroanatomy and neurophysiology of animals. Examining the cytoarchitecture and myeloarchitecture of reptiles (Dart, 1934[20]), monotremes (Shellshear, 1929; Abbie, 1940), marsupials (Abbie, 1942), non-human primates (Sanides, 1970), and more specifically macaques (Pandya and Yeterian, 1985), it can be seen that the cortex is composed of two parts which are more intra-similar than they are inter-similar in cytoarchitecture and connectivity. It has since been found in mice, rats, cats, and macaques that this general organization stems from the developing cortex being patterns by morphogens emanating from the paleocortex (now the pyriform cortex) and the archaeocortex (now the hippocampus), with the areas patterned by the paleocortex containing the ventral "what" stream and the areas patterned by the archaeocortex containing the dorsal "where/how" stream (Goulas et al., 2019).

On top of the dual origin of cortex, in primates the dorsal stream has largely superseded the superior colliculus in computing vision for motion in primates (Kaas et al., 2022). In most other vertebrates, including non-placental mammals, visual input first goes to the superior colliculus before the thalamus and then cortex (Revsin and Karten, 1967; Flandrin et al., 1981). It is then in the superior colliculus that most of the computation underlying vision is done. This changed in mammals because most of mammalian existence has been nocturnal, which lightened the evolutionary pressure

[20] A handful of lepidosaur reptiles and one turtle, whatever those even are.

placed on the mammalian visual system. While mammals evolved much more sensitive olfaction[21], audition, and somatosensation (Rowe, 1996; Rowe and Shepherd, 2016) their vision degraded, losing two kinds of cone receptor, foveae, extraocular photopigments, and a good portion of eye size (Hall et al., 2012). After the dinosaurs died[22] (Schulte et al., 2010), mammals were able to expand into crepuscular and even diurnal niches that were once dominated by the dinosaurs (Anderson and Wiens, 2017). This put evolutionary pressure back onto mammals' visual system. Many mammals grew larger eyes, increasing visual acuity. Primates even re-evolved a cone receptor, enabling our beloved color vision (Heesy and Hall, 2010). Also during this time the mammalian cortex expanded. The ancestral dorsal visual stream went from the eye to the superior colliculus, then through the pulvinar to area MT of the cortex. As the cortex grew, MT became more and more interconnected with V1, eventually becoming almost entirely driven by V1 (Rosa et al., 2000; Leopold, 2012; Kaas et al., 2022). What this all means is that the human visual system is largely bifurcate, with each visual stream evolved for a different purpose, but as of evolutionarily recently both streams are largely cortical. Cortical structures can be semantically activated.

In addition to the differences in cell composition and connectivity outlined above, the two visual streams also have different proportions of input from the magno- and parvo- cellular retinal ganglion cells. The characteristics of the different inputs were used to design two tasks, each of which was specifically suited to one visual stream or the other. The following paper presents data from two tasks which were designed to take advantage of the unique characteristics of the different inputs to the two visual streams in

[21] The giant anteater can smell ants from miles away.
[22] Rest in peace, my birdy buddies. You're never far from my thoughts.

order to argue for the hypothesis that semantic objects are perceived differently dependent on what areas across the two visual streams they are processed by. It is not just the size of a semantic object's representation that matters, but it is also where across the two visual streams that the representation is processed.

Mugs And Plants: Object Semantic Knowledge Alters Perceptual Processing With Behavioral Ramifications

Dubbelde, D., and Shomstein, S. (accepted at Psychological Science)

Introduction

Object perception underlies effective engagement with our environment. When examining a potted plant, for example, precise access to long term semantic memory may be needed to facilitate detailed identification of the plant to determine if it needs water. However, when reaching for a mug with your morning coffee, detailed identification may be less important (e.g., which one of your 10 favorite mugs it is), but computing a motor plan to bring the mug to your mouth without spilling is pertinent. The visual object processing necessary to identify plants and manipulate mugs progresses along two distinct but interacting neural processing pathways, each of which subserves different end goals (Mishkin & Ungerleider, 1982; Kravitz et al., 2011). While the hypothesis of biased processing across the two pathways is widely accepted, it is unclear whether there are behavioral ramifications that directly relate to object processing being biased to different pathways. We provide evidence that object semantic knowledge evokes differences in object perception based on the recruited regions across the visual pathways. Importantly, we gain insight into neural processing by using purely behavioral methods that are custom tailored to the response properties of the neurons in both the dorsal (higher temporal resolution) and ventral pathways (higher spatial resolution).

The ventral pathway, known as the 'what' pathway, is characterized by object feature selectivity and projects anteriorly from the occipital visual areas toward the hippocampus

and the inferior temporal cortex (Kravitz et al., 2013). The dorsal pathway, colloquially known as the 'where' or 'how' pathway, courses from the occipital visual areas into the parietal cortex. The dorsal pathway is understood to subserve object processing for the purposes of reaching, grasping, and manipulation. While both pathways contribute to visual perception, the demands of perceiving a specific object can differentially engage the two pathways such that the dorsal pathway provides greater contribution for visual processing of objects that are directly relevant for manipulation (e.g., hammer, plunger, or mug). Evidence for dynamic and biased recruitment of the visual pathways has been garnered from a wide range of techniques and paradigms. For example, while presenting participants with objects of various categories, including tools, places, animals, and faces, an increase in dorsal pathway processing, specifically in the left ventral premotor and posterior parietal cortices, is observed exclusively in response to tool presentation (Chao & Martin, 2000). Similar results evidencing a dorsal bias for manipulable objects are observed in functional magnetic resonance imaging studies (Noppeney et al., 2006a; Mahon et al., 2007, Almeida et al., 2013, Chen et al., 2018) and with various other paradigms including interocular suppression (Fang and He, 2005) and continuous flash suppression (Almeida et al., 2008; Almeida et al., 2010; but see Sakuraba et al., 2012; Almeida et al., 2014).

While neurophysiological evidence has been convincing in showing selectivity for objects across the two streams, the impact that differential processing across the two pathways has on object perception and subsequent behavior has not been characterized. Hypotheses regarding how dynamic recruitment of the visual pathways influences perception and behavior are based on an asymmetry in the cellular innervation of the two

pathways that overlaps with the separate magnocellular and parvocellular channels identified in anatomical architecture of visual processing (Maunsell et al., 1990; Baizer et al., 1991; Ferrera et al., 1992). The asymmetry in input endows each pathway with different response properties in accordance with cell stimulus preferences within the magno- and parvo-cellular channels. The magno- and parvo-cellular channels originate from different types of ganglion cells within the retina and course separately through different layers of the lateral geniculate nucleus (Leventhal et al., 1981; Perry et al., 1984) to innervate separate layers of V1 (Blasdel & Lund, 1983). From V1, the parvocellular channel can be followed into V2, V4, and then into the inferior parietal cortex while the magnocellular channel can be traced into different regions of V2, through V3d and MT, and into the posterior parietal cortices (DeYoe & van Essen, 1988; Livingstone & Hubel, 1988).

Differing innervation of the two visual pathways leads to different response properties that convey information of different spatial and temporal resolutions. The heavily myelinated magnocellular channel is derived from the parasol ganglion cells with relatively large receptive fields, spanning large regions of the retina (Maunsell et al., 1999a; Nassi & Callaway, 2009). The features of the parasol retinal ganglion cells enable the dorsally biased magnocellular channel to encode information with higher temporal resolution than the parvocellular channel (Pokorny and Smith, 1997; but see Maunsell et al., 1999b). Conversely, the parvocellular channel is derived from the midget retinal ganglion cells which receive input primarily from cone receptors and have smaller receptive fields (Nassi & Callaway, 2009). The features of the midget retinal ganglion cells enable the ventrally biased parvocellular channel to encode information with a

higher spatial resolution than the magnocellular channel (Derrington & Lennie, 1984; Leonova et al., 2003; McAnany & Alexander, 2008).

The connections between the magnocellular channel and temporal resolution and the parvocellular channel and spatial resolution have been used to study the neural basis of several behavioral effects ranging from spatial attention (Yeshurun, 2004) to the effects of fear on visual perception (Bocanegra and Zeelenberg, 2009). Most relevantly, the asymmetric contributions of the magno- and parvo-cellular channels to the dorsal and ventral pathways have been proposed as a mechanism for near-hand effects on perception. In one such study, it was hypothesized that when the participants' hands were near the stimuli then processing would be biased to the dorsal pathway, which is more predominantly magnocellular, leading to enhanced sensitivity to temporal blinks. Conversely, when the participants' hands were far from the stimuli then processing would be biased to the ventral pathway, which is more predominantly parvocellular, leading to an enhanced sensitivity to spatial gaps (Gozli et al., 2012). Consistent with differential engagement of two streams, it was observed that placing hands near the stimuli increased

participants' sensitivity to temporal blinks, while moving hands away from the screen increased sensitivity to spatial gaps.

Knowing that semantic knowledge of manipulability biases processing to the dorsal stream, and that the dorsal and ventral streams have different response profiles due to differential innervation by the magnocellular and parvocellular channels, we hypothesize that semantic knowledge of object manipulability should have consequences for perceptual processing. Thus, because semantic knowledge of an object's

manipulability determines which pathway will be biased, we predict that (1) manipulable objects, that elicit a higher degree of magnocellularly biased dorsal processing, are processed with higher temporal resolution and (2) non-manipulable objects, that rely more on parvocellularly biased ventral processing, are processed with higher spatial resolution. Across five experiments, this hypothesis is tested by comparing spatial and temporal resolution across two object groups: manipulable and non-manipulable objects.

Method

Participants

All participants were recruited from The George Washington University participant pool, gave informed consent according to the George Washington University's institutional review board (IRB), were naïve to the purpose of the experiment, and all reported normal or corrected-to-normal vision. In experiment 4, which utilized color stimuli no participant reported colorblindness.

For each experiment sample sizes were chosen based on behavioral studies demonstrating similar effects (i.e., Gozli et al., 2012). For experiment 1, 26 undergraduate students (19 female; average age: 19.2; 6 left-handed) were recruited. No participant was removed from the analyses. For experiment 2, 27 undergraduate students were recruited. Twenty-six participants' data (14 female; average age: 19.0; 1 left-handed) were analyzed after one participant was cut for below chance accuracy in one of the conditions. For experiment 3, 33 undergraduate students were recruited. Thirty participants' data (26 female; average age: 19.07, 3 left-handed) were analyzed after three participants were cut for below chance accuracy in at least one of the conditions. For experiment 4, 40 undergraduate students were recruited. Twenty-five participants' data

(14 female; average age: 19.0; 1 left-handed) were analyzed after six participants were cut for not finishing the experiment and nine participants were cut for having less than chance accuracy in at least one of the conditions. For experiment 5, 26 undergraduate students were recruited. Twenty-five participants' data (18 female; average age: 19.5, 0 left-handed) were analyzed after one participant was cut for chance accuracy in all conditions.

Apparatus and Stimuli

All experiments were presented on a 24" Acer GN246 HL monitor with a refresh rate of 144 Hz, positioned at a distance of 60 cm from the viewer in a dark room. The experiment was generated and presented using PsychoPy v1.82.

The stimuli in Experiments 1, 2, 4, and 5 consisted of line drawings of real-world, everyday objects drawn both from The Noun Project, an online repository of object icons and clip art (https://thenounproject.com). The object stimuli consisted of 10 line-drawings of manipulable objects, and 10 line-drawings of non-manipulable objects. The manipulable objects were as follows: snow shovel, handsaw, plunger, screwdriver, hammer, wrench, knife, bottle opener, spatula, and mug. The non-manipulable objects were: fire hydrant, picture frame, window, toilet, candle, garbage can, water fountain, potted plant, fan, and lamp. The stimuli were displayed as large as possible in a 4° x 4° area and all stimuli were presented in black (HSV = 0, 0, 0) on a dark gray background (HSV = 0, 0, 50). All objects are displayed in Figure 1.

Objects were controlled for low-level differences by calculating the mean luminance, size, aspect ratio (i.e., measure of elongation), and spatial frequency (average distance

from origin of the points calculated from a 2D fast Fourier transform) for each object. Assessed with independent samples t-tests, there were no mean differences between object groups in luminance($t(18) = -0.960$, $p = 0.350$), mean size (t-test ($t(18) = 1.043$, $p = 0.311$), aspect ratio ($t(18) = 1.209$, $p = 0.242$), or spatial frequency (t-test ($t(18) = -0.155$, $p = 0.879$). The width of the bottom line was controlled for each object on which the gap appeared. There was no significant difference between manipulable and non-manipulable objects in bottom line width in an independent samples t-test ($t(18) = 0.922$, $p = 0.369$).

For Experiment 3, the stimuli consisted of images of real-world, everyday objects drawn from Google image search and manipulated in GIMP 2.10.4. The object stimuli consisted of 10 images of the same manipulable objects, and 10 images of the same non-manipulable objects used in Experiments 1,2, 4, and 5. The stimuli were displayed as large as possible in a 4° x 4° area and all stimuli were presented in color on a dark gray background (HSV = 0, 0, 50). All objects are displayed in Figure 1.

Additionally, the colored objects used in Experiment 3 were similarly controlled for low-level differences by calculating the mean size, hue, saturation, value, aspect ratio, and spatial frequency. There was no significant differences between manipulable and non-manipulable objects in an independent samples t-test for mean size ($t(18) = -1.027$, $p = 0.318$)hue ($t(18) = -0.309$, $p = 0.761$), saturation ($t(18) = -0.031$, $p = 0.976$), value ($t(18) = 0.032$, $p = 0.975$), aspect ratio ($t(18) = -0.094$, $p = 0.927$), or spatial frequency ($t(18) = -1.466$, $p = 0.160$) Additionally, we controlled the width of the bottom line of each object on which the gap would appear. There was no significant difference between

manipulable and non-manipulable objects in bottom line width in an independent samples t-test ($t(18) = 0.557$, $p = 0.584$).

A

Fire Hydrant	Picture Frame	Window	Toilet	Candle	Garbage Can	Water Fountain	Potted Plant	Fan	Lamp
0.010°	0.006°	0.008°	0.006°	0.010°	0.005°	0.018°	0.013°	0.006°	0.007°
Snow Shovel	**Handsaw**	**Plunger**	**Screw-driver**	**Hammer**	**Wrench**	**Knife**	**Bottle Opener**	**Spatula**	**Mug**
0.007°	0.011°	0.007°	0.012°	0.010°	0.007°	0.010°	0.008°	0.015°	0.010°

B

Fire Hydrant	Picture Frame	Window	Toilet	Candle	Garbage Can	Water Fountain	Potted Plant	Fan	Lamp
0.010°	0.013°	0.012°	0.009°	0.009°	0.007°	0.011°	0.012°	0.007°	0.008°
Snow Shovel	**Handsaw**	**Plunger**	**Screw-driver**	**Hammer**	**Wrench**	**Knife**	**Bottle Opener**	**Spatula**	**Mug**
0.016°	0.018°	0.007°	0.014°	0.008°	0.009°	0.009°	0.009°	0.009°	0.013°

Figure 1 *Stimulus set used in the experimental paradigms. (A) Line drawings of manipulable and non-manipulable objects used in Exp. 1. (B) Pictures of real-world manipulable and non-manipulable objects used in Experiment 3. The numbers beneath each set of objects correspond to the size of the spatial gap titrated to each object.*

Task

In Experiment 1, each trial began with a single central fixation point, which subtended a visual angle of 1° x 1°. The central fixation point was rendered in white (HSV = 0, 0, 100). After 1000 ms of fixation presentation, a single object line drawing would appear 4° to the left or right of the fixation point. The side of presentation was counterbalanced such that each participant saw an equal number of stimuli on the left and on the right. The stimuli were displayed as large as possible within a 4° x 4° area and all

stimuli were presented in color on a dark gray background (HSV = 0, 0, 50). All objects are displayed in Figure 1.

In one half of the experiment, the object appeared with or without a spatial gap in the center of the bottom line of the object. The objects had an equal probability of having a spatial gap or not having a spatial gap. The objects appeared for 100 ms and participants were to report the presence of the spatial gap by pressing the right control button (present) or the left control button (absent) on the keyboard. Feedback was presented on incorrect trials only.

A staircase procedure was used to calibrate the size of the gap to each object to ensure that the gap was equally perceptible across each object regardless of each objects' individual characteristics (Pelli and Bex, 2013). Sixteen undergraduate students (13 female; average age: 18.8; 4 left-handed) from the George Washington University participated in exchange for course credit. In each trial the object would be presented with or without the spatial gap which would begin with a size of 0.025° visual angle. If the gap was detected correctly for two consecutive trials, the gap decreased by one 0.005° step. If the gap was missed, or a false alarm was made to the absence of a gap, for two consecutive trials, the gap was increased by one 0.005° step. The gap was calibrated until twenty trials had been completed or until the staircase reversed direction (2 correct followed by 2 incorrect or vice versa) three times. The gap size for each object is displayed in figure 1.

In the other half of the experiment, the object appeared with or without a temporal blink. The objects had an equal probability of having a temporal blink or not having a temporal blink. The blinks were 16 ms long. The object first appeared for 96 ms, followed by the temporal blink, and the object then appeared a second time for 32 ms. Participants

were to report the presence of the temporal blink by pressing the right control button and absence of the temporal blink by pressing the left control button on the keyboard. Feedback was presented on incorrect trials only.

The presentation of stimulus type was counterbalanced across participants such that half of the participants were first presented with spatial gaps and half of the participants were presented with temporal blinks first. Participants completed a total of 720 trials broken into 6 blocks, 360 trials of spatial gaps and 360 trials of temporal blinks.

Experiment 2 was identical to Experiment 1 except each stimulus was presented upside down. The spatial gap was presented in the same place as it was in experiment 1 relative to the object.

Experiment 3 was identical to experiment 1, except more realistic images were used instead of line drawings. Similarly to the line drawing stimuli, a staircase procedure was used to calibrate the size of the gap to each object to ensure that the gap would be equally perceptible across each object regardless of each objects' individual characteristics (Pelli and Bex, 2013). Twenty-three undergraduate students (18 female; average age: 18.9; 4 left-handed) from the George Washington University participated in exchange for course credit and the same procedure from Experiment 1 was used. The gap size for each object is displayed in figure 1.

Experiment 4 was similar in trial structure to the temporal blink condition of Experiment 1, except instead of temporal blink detection each trial had a temporal blink that was either short or long in duration. The objects had an equal probability of having a short or a long temporal blink. In the short blink condition, the object first appeared for 96 ms, followed by a 16 ms temporal blink, and then the object appeared a second time

for 32 ms. In the long blink condition, the object first appeared for 64 ms, followed by a 48 ms temporal blink, and then the object was presented for a second time for 32 ms. Participants were to report a presence of a short temporal blink by pressing the 'c' key and a long temporal blink by pressing the 'm' key on the keyboard. Feedback was presented exclusively on incorrect trials. Participants completed a total of 720 trials broken into 6 blocks, 360 trials of short blink duration and 360 trials of long blink duration.

Experiment 5 was identical to the spatial gap procedure of Experiment 1 except the background color was either green (HSV = 110, 30, 90) or red (HSV = 5, 30, 90). The background color was counterbalanced across participants such that half of the participants would first be presented with the green background and half of the participants would first be presented with the red background. Participants completed a total of 720 trials broken into 6 blocks, 360 trials of green backgrounds and 360 trials of red backgrounds.

Results

Experiment 1

The spatial and temporal paradigms used d' as a measure of perceptual sensitivity (Fig. 2). A two-way repeated-measures analysis of variance (ANOVA) was conducted on d' with object group (manipulable, non-manipulable) and stimulus type (gap, blink) as within-subject variables (Fig. 2C, left). The ANOVA revealed no significant main effect of object group ($F(1, 25) = 2.342$, $p = 0.138$, $\eta_p^2 = 0.086$), and a significant main effect of task type ($F(1, 25) = 50.953$, $p < 0.001$, $\eta_p^2 = 0.671$) such that d' sensitivity was higher for blinks than for gaps ($M_{gaps} = 1.926$, 95% CI [1.673, 2.179]; $M_{blinks} = 2.967$, 95% CI [2.680, 3.254]). Importantly, and consistent with the hypothesis of differential

engagement of two pathways depending on object utility, a significant two-way interaction between object group and task type ($F(1, 25) = 6.772$, $p = 0.015$, $\eta_p^2 = 0.213$) was observed, driven by the difference between object groups in the gap condition ($F(1, 25) = 9.888$, $p = 0.004$, $\eta_p^2 = 0.283$), with non-manipulable objects having a higher d' for gaps than manipulable objects ($M_{non\text{-}manipulable} = 2.021$, 95% CI [1.756, 2.285]; $M_{manipulable} = 1.831$, 95% CI [1.589, 2.073]). The interaction effect and the driving simple effect are consistent with the prediction that non-manipulable objects, given their higher reliance on the ventral pathway, should yield higher sensitivity in the detection of spatial gaps than manipulable objects. Notably, the expected higher sensitivity to temporal gaps in the manipulable object set was not supported. The possibility of an advantage for manipulable objects in temporal sensitivity is further examined in experiment 4.

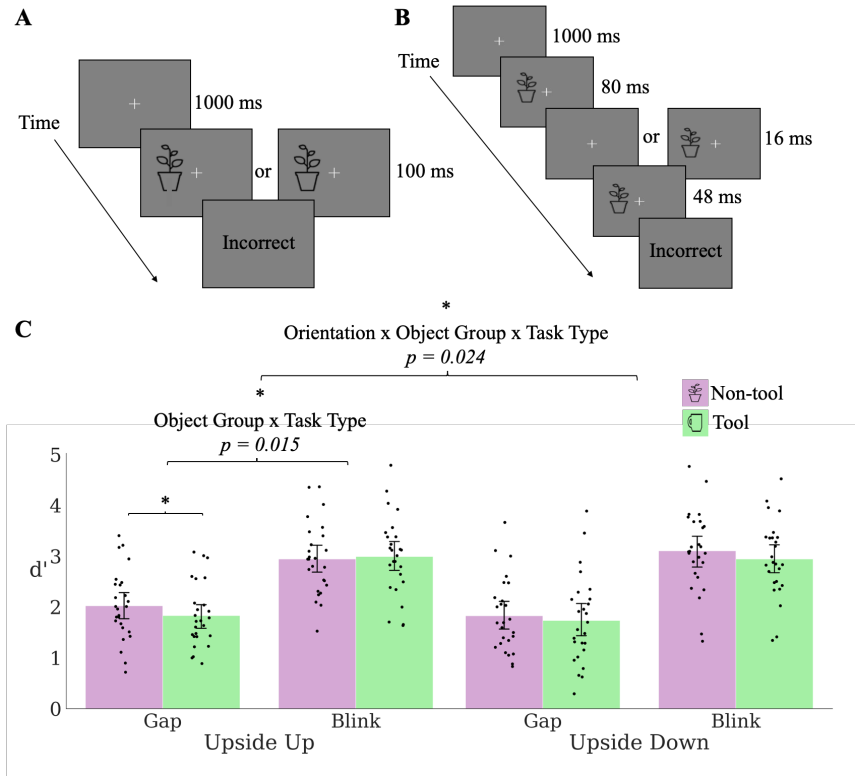


Figure 2. *The paradigm and results for Experiments 1 and 2. Participants maintained fixation on the center cross. An object appeared to the left or right of fixation. A) For the spatial-gap paradigm, participants indicated by keypress if the bottom line of the presented object contained a gap. B) For the temporal paradigm, participants were to indicate whether the object flickered. C, left) The results from Experiment 1 in which the object stimuli were presented upside up. C, right) The results from Experiment 2 in which the objects were presented upside down.*

Experiment 2

To further probe the hypothesis that semantic knowledge of object utility (manipulable or non-manipulable) biases the pathway that will ultimately process the object, we designed Experiment 2 to demonstrate that the hypothesized pathway biasing will not occur if the semantic identity of the objects is obscured (Firestone and Scholl, 2016). The same paradigm and objects from Experiment 1 were used but the objects were inverted (upside-down). The inversion preserved the low-level features of each object while impairing participants' ability to rapidly recognize and access their semantic knowledge of the objects' function. Inversion has been found to interfere with recognition of faces and objects (Diamond and Carey, 1986). Following from the hypothesis that semantic knowledge of object's utility drives the perceptual difference between manipulable and non-manipulable objects, we predict that object inversion should reduce the difference between manipulable and non-manipulable objects. A two-way repeated-measures ANOVA was conducted on d' with object group (manipulable, non-manipulable) and task type (gap, blink) as within-subject variables. The ANOVA

revealed a significant main effect of object group ($F(1, 25) = 5.860$, $p = 0.023$, $\eta_p^2 = 0.190$) such that d' was higher for non-manipulable objects than for manipulable objects ($M_{non\text{-}manipulable} = 2.466$, 95% CI [2.174, 2.758]; $M_{manipulable} = 2.340$, 95% CI [2.029, 2.650]) (Fig. 2C, right). There was also a significant main effect of task type ($F(1, 25) = 82.298$, $p < 0.001$, $\eta_p^2 = 0.767$) such that d' was higher for blinks than for gaps ($M_{gaps} = 1.782$, 95% CI [1.476, 2.088]; $M_{blinks} = 3.023$, 95% CI [2.723, 3.320]). As predicted, there was no significant interaction between object group and task type ($F(1, 25) = 0.508$, $p = 0.483$, $\eta_p^2 = 0.020$).

To assess the non-significant interaction effect, a Bayes factor analysis was conducted using the Bayesian repeated measures ANOVA in JASP (van den Bergh et al., 2019) comparing the posterior probability of a model with the main effects of stimulus type and object group but no interaction term as the null hypothesis (H_0) to the posterior probability of a full model with the main effects and the interaction term as the alternative hypothesis (H_1). The Bayes factor analysis evaluating the non-significant interaction effect yielded a Bayes factor of 4.83 ($BF_{01} = 4.83$), indicating substantial evidence for the null hypothesis which does not include the interaction (Kass & Raftery, 1995).

Lastly, in order to statistically demonstrate that results of the inversion experiment are indeed different from those observed in Experiment 1, data were subjected to a between-subject ANOVA with object orientation (upright, inverted) as a between subjects variable and object group and task type as within subject factors. The ANOVA revealed a significant three-way interaction between object group, task type, and

orientation ($F(1, 50) = 5.381$, $p = 0.024$, $\eta_p^2 = 0.097$) such that, as predicted, orientation significantly reduced the effect for the inverted objects.

Experiment 3

While Experiment 1 provides evidence for biased engagement of the ventral pathway for non-manipulable object perception in a spatial task, it could be argued that despite careful low-level feature controls (e.g., luminance, size, elongation, spatial frequency) an uncontrolled low-level difference between manipulable and non-manipulable objects is responsible for driving the manipulable vs. non-manipulable advantage in the spatial gap task.

Experiment 3 used the same paradigm as Experiment 1 except the line drawings were replaced by real-world images of corresponding objects, such that a line drawing of a candle was replaced with a picture of a candle, etc. (Fig. 1B). The prediction remained the same as in the original experiment. If object semantic knowledge determines which visual pathway object processing is biased to, then non-manipulable objects will be biased to the ventral pathway, leading to higher sensitivity (as measured by d') in the spatial gap task. This would replicate the pattern of performance observed in Experiment 1. A two-way repeated-measures ANOVA was conducted on d' with object group (manipulable, non-manipulable) and task type (gap, blink) as within-subject variables. The ANOVA revealed a significant main effect of object group ($F(1, 29) = 28.211$, $p < 0.001$, $\eta_p^2 = 0.493$) with non-manipulable objects having a higher sensitivity than manipulable objects ($M_{\text{non-manipulable}} = 2.361$, 95% CI [2.054, 2.668]; $M_{\text{manipulable}} = 2.047$, 95% CI [1.773, 2.320]) (Fig. 3). There was also a significant main effect of task type

($F(1, 29) = 8.413$, $p = 0.007$, $\eta_p^2 = 0.225$) with blinks having a higher average sensitivity than gaps ($M_{gaps} = 1.998$, 95% CI [1.758, 2.238]; $M_{blinks} = 2.410$, 95% CI [2.068, 2.751]). Importantly, a two-way interaction between object group and task type was also significant ($F(1, 29) = 13.061$, $p = 0.001$, $\eta_p^2 = 0.311$), driven by a simple main effect in the gap condition ($F(1, 29) = 46.424$, $p < 0.001$, $\eta_p^2 = 0.616$) such that non-manipulable objects had a higher d' for gaps than did the manipulable objects ($M_{non\text{-}manipulable} = 2.238$, 95% CI [1.982, 2.494]; $M_{manipulable} = 1.758$, 95% CI [1.535, 1.981]). In addition to serving as a low-level control and a further test of the hypothesis, these results replicate Experiment 1, providing strong additional support for the hypothesis that the perceptual differences between manipulable objects and non-manipulable objects is due to the semantic content of the objects.

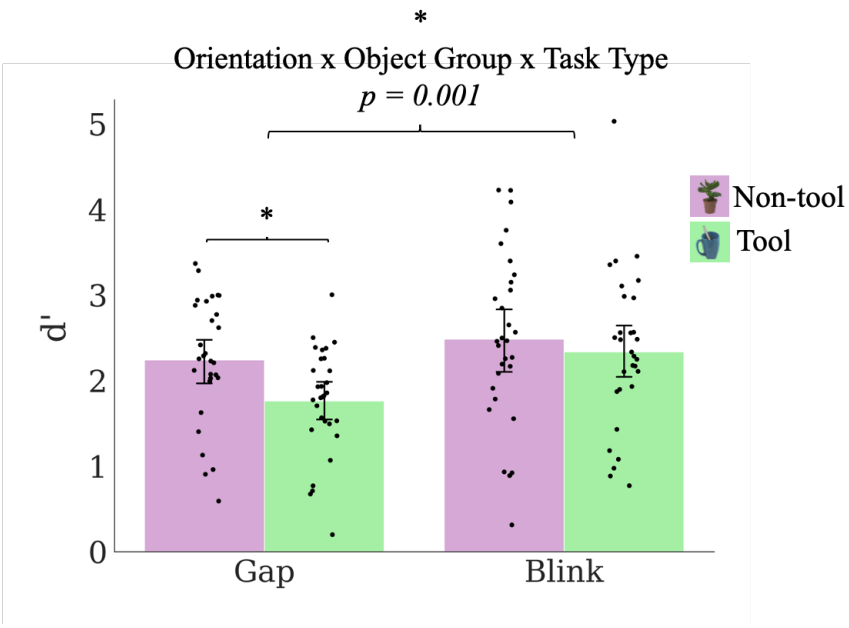


Figure 3. *Experiment 3 results: a control manipulation of Experiment 1 in which real object images were used instead of line drawings.*

Experiment 4

While the first three experiments yielded strong supporting evidence that gap detection is better on non-manipulable objects than manipulable objects, perhaps resulting from bias toward the ventral pathway, manipulable objects failed to elicit an advantage in temporal resolution, failing to provide evidence of a dorsal pathway bias for manipulable objects. One explanation for the null effect is that the temporal gap detection task could also be construed as an abrupt onset detection task due to the object suddenly reappearing after the temporal blink. Abrupt onsets are known to be highly salient and easily detectable (Yantis and Jonides, 1984) and evoke strong activity in the lateral intraparietal area, a known attentional area with strong connections to the superior colliculus (Kusunoki et al., 2000). The strong connection between the LIP and the superior colliculus provides a mechanism by which abrupt onsets may bypass the parvo- and magno-cellular two visual pathways framework which we aimed to test, thereby collapsing any differences in the d' observed for manipulable and non-manipulable objects. Another possible explanation for the null effect could be that the temporal gap detection as was too easy (ceiling effect), evidenced by high d' values in the temporal task. In order to have a task that is more difficult and is more specifically targeted to the high temporal resolution of the magnocellular channel, the blink paradigm used in Experiments 1 through 3 was re-designed to a discrimination task[23]. Objects were presented for 80 ms, removed from the screen for either a 16 ms or 48 ms blink, and redisplayed for 48 ms (Fig. 4). Participants' task was to indicate whether the blink duration was short or long, with d' being calculated with short blinks being considered

[23] We thank Ed Awh, from the University of Chicago, for the helpful suggestion.

hits or misses and long blinks considered as correct rejections or false alarms. The prediction was that manipulable objects should have a higher d' than non-manipulable objects due to the increased magnocellular input, and therefore temporal resolution, of the dorsal pathway. A t-test was used to analyze the difference between manipulable and non-manipulable objects in d' ($M_{non\text{-}manipulable}$ = 1.562, 95% CI [1.354, 1.770]; $M_{manipulable}$ = 1.620, 95% CI [1.385, 1.855]), with no significant effect being found (t(48) = -0.362, p = 0.719).

Due to the temporal nature of the task, RT was also analyzed, reasoning that the any benefit in temporal resolution for manipulable objects may manifest as a speeded reaction time rather than a sensitivity benefit. A two-way repeated-measures ANOVA was conducted on RT with object group (manipulable, non-manipulable) and blink duration (short, long) as within-subject variables. Only the RTs for correct responses were analyzed. The ANOVA revealed no significant main effect of object group (F(1, 24) = 4.080, p = 0.055, η_p^2 = 0.145) (Fig. 4). There was a significant main effect of blink duration (F(1, 24) = 4.850 p = 0.038 η_p^2 = 0.168) such that the short duration elicited a higher RTs than did the long duration (M_{short} = 642.680, 95% CI [604.400, 680.960]; M_{long} = 621.997, 95% CI [581.794, 662.200]). The two-way interaction between object group and blink duration was significant (F(1, 24) = 9.011, p = 0.006, η_p^2 = 0.273) driven by a simple main effect in the short blink condition (F(1, 24) = 11.710, p = 0.002, η_p^2 = 0.328) such that manipulable objects had a lower RT for short blinks than did the non-manipulable objects ($M_{non\text{-}manipulable}$ = 652.210, 95% CI [612.360, 692.060]; $M_{manipulable}$ = 633.150, 95% CI [596.441, 669.859]). In accuracy, no effect was observed for object group (F(1, 24) = 3.666, p = 0.068, η_p^2 = 0.133), blink duration (F(1, 24) = 1.509, p =

0.231, $\eta_p^2 = 0.059$), nor the interaction ($F(1, 24) = 2.000$, $p = 0.170$, $\eta_p^2 = 0.077$), suggesting that all significant effects were absorbed by the response time measure and no speed-accuracy tradeoffs were observed, indicating consistency with response time results.

These results, specifically the simple main effect in the short gap condition with manipulable objects having a lower RT than non-manipulable objects, lend some support to our hypothesis that manipulable objects would elicit higher temporal resolution than non-manipulable object processing. While this result is not as strong as the observed benefit for non-manipulable objects in spatial resolution, nor is it manifested in sensitivity, but it does suggest that an advantage in temporal resolution is indeed present for manipulable objects. Understanding the exact nature of this temporal resolution advantage may be a fruitful direction for future studies to take.

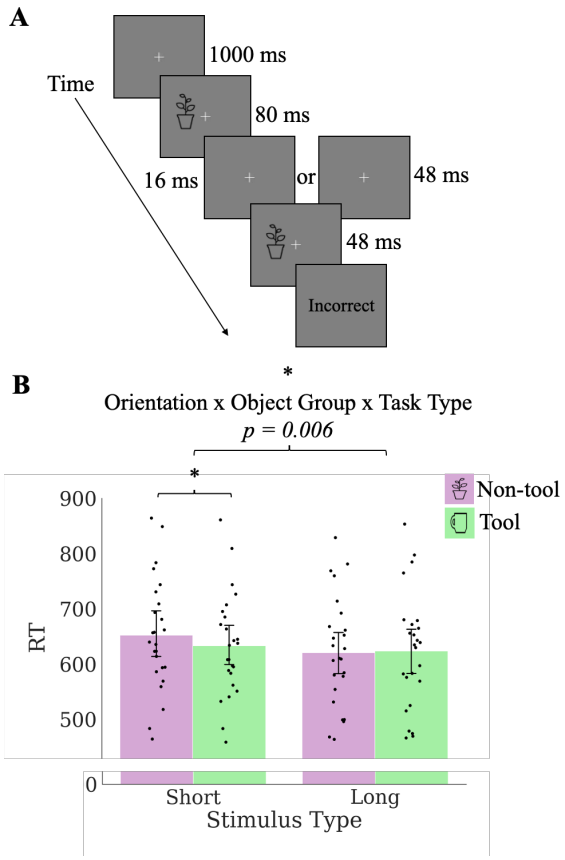


Figure. 4. *The paradigm and results for experiment 4. A) Participants maintained fixation on the center cross. An object would then appear to the left or right of fixation. The objects would flicker by being removed from the screen for 16 ms or 48 ms. Participants were to indicate by keypress if the object flicker duration was short or long. B) The results from experiment 4.*

Experiment 5

The last test of our hypotheses is derived from the neurophysiological differences between the dorsal and ventral pathways. It was reasoned that if the differences in spatial gap sensitivity for non-manipulable objects are mechanistically derived from the magnocellular and parvocellular dichotomy of the two visual pathways, then the effect

should be modulated through manipulation of the processing within the pathways. Ambient red light has been shown to suppress activity in the magnocellular channel because of the large number of visually responsive cells with red inhibitory surrounds in their receptive fields (Wiesel & Hubel, 1966) and has been used to demonstrate the contribution of magnocellular processing in other behavioral effects such as fear processing and near hand effects (West et al., 2010; Abrams and Weidler, 2014). To test our hypothesis that the spatial resolution difference between manipulable and non-manipulable objects that we observed in experiments one and three is due to the differential input of the magnocellular channel to the two visual streams, the spatial resolution paradigm from Experiment 1 was used with the manipulation of background color varying between green or red (Fig. 5A). Due to the suppression of the magnocellular channel by red light, it was predicted that the color of the background should modulate the perceptual difference in spatial gap detected observed in Experiment 1. A two-way repeated-measures ANOVA was conducted on d' with object group (manipulable, non-manipulable) and background color (green, red) as within-subject variables. The ANOVA revealed a significant main effect of object group ($F(1, 25) = 17.171$, $p < 0.001$, $\eta_p^2 = 0.407$) with non-manipulable objects having a higher average d' than manipulable objects ($M_{\text{non-manipulable}} = 2.250$, 95% CI [1.941, 2.559]; $M_{\text{manipulable}} = 2.022$, 95% CI [1.757, 2.287]) (Fig. 5B). Note that this main effect per se is a second replication of the spatial effect seen in Experiments 1 and 3. There was no significant main effect of background color ($F(1, 25) = 2.138$, $p = 0.156$, $\eta_p^2 = 0.079$) but, as predicted, there was a significant two-way interaction between object group and background color ($F(1, 25) = 4.444$, $p = 0.045$, $\eta_p^2 = 0.151$) such that the effect was

increased with the red background. It was hypothesized that if the perceptual differences between manipulable and non-manipulable objects are mechanistically derived from the magno- and parvo-cellular processing in the two visual pathways, then suppression of the magnocellular processing with red light should modulate the effect. The results of Experiment 5 support our hypothesis by demonstrating an increase of the perceptual difference between manipulable and non-manipulable objects with red light.

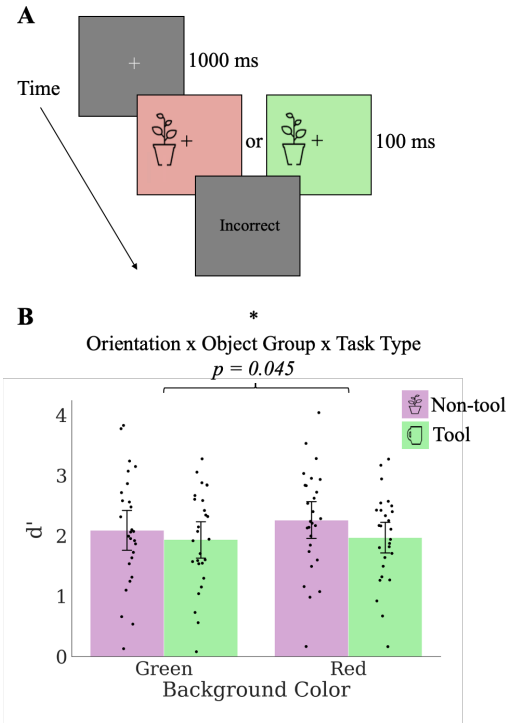


Figure 5. *The paradigm and results of experiment 5. (A) Participants maintained fixation on the center cross. An object then appeared to the left or right of fixation. As in Experiment 1, participants were asked to detect the presence of small gap in the bottom outline of the object. The background color (red or green) was manipulated. (B) The results from experiment 5.*

Discussion

We hypothesized that manipulable and non-manipulable objects, due to differential recruitment of the visual pathways, would elicit perceptual differences stemming from the characteristic spatial and temporal resolutions associated with the magno- and parvo-cellular inputs. In five experiments we provide strong evidence in support of behavioral consequences driven by physiological and anatomical differences of the two visual pathways, and argue that semantic knowledge of object manipulability guides processing along a particular pathway. If an object has a strong action association, processing is largely determined by activity in the dorsal pathway which is not found with objects that lack strong action association (Chao & Martin, 2000; Noppeney et al., 2006; Mahon et al., 2007). The increased level of activity in the parietal regions endows the perception of action-associated objects with greater access to the magnocellular channel which preferentially courses through the dorsal pathway. Without the enhanced dorsal pathway activity evoked by action associations, object processing is more dependent on the ventral pathway, which has a higher ratio of parvocellular channel input than does the dorsal pathway (Maunsell et al., 1990; Baizer et al., 1991; Ferrera et al., 1992). Based on differential magno- and parvo-cellular input to the two pathways, we predicted that the perception of objects with strong action associations results in an increased temporal resolution and the perception of objects without such associations results in an increased spatial resolution. In the experiments reported here, we offer evidence that objects are perceived with different spatial and temporal resolutions depending on object semantic knowledge of manipulability. In two follow-up control experiments, further evidence was provided in support of the hypothesis that the

difference between manipulable and non-manipulable objects in the spatial task was driven by the semantic knowledge of the objects rather than possible low-level visual features. Namely, it was observed that the effect is curtailed by impeding the access of object semantic knowledge through inversion and that the perceptual differences between manipulable and non-manipulable objects replicates with a separate set of more realistic object images (with different low-level properties). Lastly, to test if the differing proportions of magno- and parvo-cellular input are responsible for the perceptual differences that were observed in our first four experiments, ambient red light was used to suppress activity of the magnocellular channel. It was observed that this increased the perceptual difference in the spatial task between manipulable and non-manipulable objects.

Based on the evidence provided, we argue that semantic knowledge of object manipulability, as defined by strong associations with an action appropriate for the item, generates the perceptual differences between manipulable and non-manipulable objects. Previous studies have shown that manipulable objects evoke a larger degree of dorsal pathway activity than do non-manipulable objects (Chao & Martin, 2000; Almeida et al., 2013). However, the origin of this difference and consequent perceptual differences between manipulable and non-manipulable objects is poorly understood.

We propose that the differences we observe between manipulable and non-manipulable objects derive directly from the near-hand effect reported by Gozli et al. (2012) in which stimuli presented proximally to the participants' hands evoked a benefit in spatial resolution, similar to the non-manipulable objects used in the experiments presented here, and stimuli presented distally from the participants' hands evoked a

benefit in temporal resolution. A possible mechanism supporting a connection between the near-hand effect and the manipulable vs. non-manipulable effect presented here derives from the bimodal cells in the anterior parietal cortex (Graziano & Gross, 1993). The bimodal cells have a somatosensory receptive field covering a part of the hand (a property possibly underlying the near-hand effect) and a visual receptive field corresponding to the visual field near the associated hand area. When an object is shown, if that object does not have a strong action association then the activity of the bimodal cells is unaffected. After the organism has learned to manually manipulate the object, the bimodal cells respond to visual presentation of the object even without somatosensory input (Zhou & Fuster, 2000). Thus, the bimodal cells, with their responsiveness to hand location and their ability to become visually activated, may be a driver of the dorsal pathway's object selectivity (Vaziri-Pashkam & Xu, 2017; Kastner et al., 2017; Freud et al., 2016), leading to the dorsal pathway bias evoked by manipulable objects (Chao & Martin, 2000). The object-specific neural activity in the dorsal pathway, divorced from its need for somatosensory input, and largely derived from magnocellular input, could then be read out during an identification task, leading to heightened temporal precision for manipulable objects. Taken together, our results demonstrate that object semantic knowledge determines the processing bias of the object and evokes subsequent behavioral repercussions for perception and for action. This finding, in conjunction with the finding that manipulability has an influence on how attention interacts with object perception (Gomez et al., 2018), may point to manipulability, supported by a detailed neural mechanism, as an exemplar of cognitive penetrability. Additionally, our work underscores the need for careful consideration of object semantic knowledge, and its

subsequent possible bias to either dorsal or ventral pathway, when object images are used not only in psychological research, but in applied settings such as display and product designs, environmental design, and in the design of various cognitive assistants.

CHAPTER 4 – ATTENTION AND PLATYPUSES

Lastly, attention. Think about attention for a second. It is such a familiar term. It is something you pay. It is something you can hold. It's something that you can lose. This intuition led William James to make the claim that "Everyone knows what attention is[24]" (James, 1890). As with everything in the brain though, it is almost infinitely more complicated. Compare our knowledge of attention to our knowledge of memory, another such familiar intuition. Most people now who have even a passing interest in the workings of the brain realize that memory is not a unitary thing, but rather a group of things in one trench coat[25]. The most obvious distinction is between short-term and long-term memory, as anyone who has aced a multiple guess exam can attest to. Studies of patient HM demonstrated that this was a neurally relevant distinction as HM had short-term memory but no way to create new long-term memory (Scoville and Milner, 1957). The number and nature of things inside the trench coat of memory is still debated, but the realization of the trench coat nature of the term 'memory' has created a lasting paradigm.

Attention is undergoing the same process, but it is several steps behind and is going more slowly. This realization of the trench coat, and the absolute adoration of metaphor, has left the study of attention in a bit of a pickle, scrambling for definition. This has even resulted in there being papers with titles like "No one knows what attention is" (Hommel et al., 2019). In the last paper presented here, we argue that sure, we may not have a full understanding of what attention is, but by clearly defining what we mean

[24] "It is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies a withdrawal from some things in order to deal effectively with others."
[25] Like a child on another child's shoulders and wearing a trench coat to appear as an adult.

by attention we may be able to sidestep the confusion. Attention isn't undefined, but maybe we should be a bit more explicit about precisely what we mean.

We start by examining prevalent model for grouping attentional effects, which includes groups for bottom-up effects (e.g. abrupt onsets, see Yantis and Jonides, 1984), top-down effects (e.g. object-based attention, see Egly et al., 1994), and hysteresis (e.g. search history matters, see Awh et al., 2012). Beyond bottom-up, top-down, and hysteresis effects though there are a large number of effects, which we as a community commonly call attentional, which do not fit neatly into the three group model in any way which doesn't severely strain the utility of such grouping. For example, it has been shown that task-irrelevant semantic objects influence the allocation of attention (Malcolm et al., 2016; Collegio et al., 2019). The task-irrelevant objects are not bottom-up as there is nothing especially salient about them in the same way that a red flashing light is salient. They are not top-down as they are completely irrelevant to doing the task and participants are never instructed to identify the objects. Similarly, they do not fit into hysteresis effects because the task-irrelevant objects in each trial are completely unrelated to the objects of the previous trial. We call uncategorizable attentional effects like this effect of task-irrelevant objects 'attentional platypuses' for two reasons. One, the attention research community absolutely loves metaphors[26]. Two, like platypuses, attentional effects like this challenge our taxonomies. In a simple taxonomy of the animal kingdom, such as the one that I was taught in grade school, there were birds and there were

[26] Also, me too.

mammals[27]. The platypus, with its fur and lactation but also its bill and eggs don't fit in this grouping at all[28].

There are two ways of responding to a challenge to your taxonomic system. Either you defend it[29], which was what evolution did for animal taxonomy, or you scrap it. We proposed scrapping the attentional taxonomy by defining attention as the functioning of dynamic priority maps. Priority maps are representations of space, similar to the retinotopic maps of early visual cortex, which are not only composed of some bottom-up information like space but also the top-down intentions of the organism (Bisley and Goldberg, 2010; Ptak, 2012; Shomstein and Gottlieb, 2016; Todd and Manaligod, 2017). By viewing attentional effects through this lens and describing them in terms of how they affect the priority maps, rather than in terms of how similar or different they are, we are putting our knowledge of attention into a common space which enables us to better understand how the factors which influence the functioning of attention compare and interact. An analogy can be drawn to how the concept of natural selection enabled us to better compare organisms to form a better, more complete understanding of the functioning of the natural world.

This utility of thinking of attention in terms of dynamic priority maps rather than box plots is much more thoroughly in the following paper. There is a question posed at the end of the paper, though which is unresolved: is defining attention as the functioning

[27] And of course fish, mollusks, crustaceans, insects and reptiles, but those are less relevant right now.
[28] Really the platypuses aren't anywhere near being birds with their last common ancestor with birds being the same as ours, somewhere about 320 million years ago when the amniotes split into synapsids (us and platypuses) and sauropsids (birds and much more importantly, dinosaurs).
[29] And tweak it. The Linnaean system has changed considerably but the basic structure is still intact.

of a dynamic priority map a falsifiable hypothesis? I'd like to add two points to this discussion.

One, "attention is the functioning of a dynamic priority map" is not a hypothesis but rather a definition. It is not a description of a natural process which attempts to capture the nature of the process, but rather it is more akin to pointing at a creature and saying 'cat.' As such, the question of falsifiability is moot[30]. We are simply trying to be more rigorous in communicating precisely what we are talking about.

Two, "attention is the functioning of a dynamic priority map" does enable us to be more rigorous if we ground the dynamic priority maps in neural circuits. We can then say if a process is attentional or not dependent on the involvement of the associated neural circuits. The computation of priority maps is attributed to several regions of the parietal cortex which are sensitive to both sensory activity and intention, upregulating the activity associated to spatial regions via feedback connections (Corbetta et al., 2000; Indovina and Macaluso, 2007; Kravitz et al., 2011). By assessing the involvement in the parietal, priority-associated circuits we can then determine if a process fits our definition of 'attentional,' enabling us to separate attentional processes from any other process which reduces the amount of information in the neural circuits. For example, the size of a visual neuron's receptive field[31] determines the amount of visual information which is lost when encoding visual stimuli in that receptive field into neuronal firing. In essence, a larger receptive field size means the neuron is computing summary statistics (i.e. averaging) over a larger amount of information. A perfect example of this can be found in the retina.

[30] It's like a cow's opinion, "moo."

[31] A receptive field is the area of the visual scene to which a neuron responds. These have traditionally been found with electrophysiological recordings of cells in anesthetized animals by moving various objects in front of the animal's eyes, A Clockwork Orange style.

The cones of the retina are very densely packed in the fovea, but from just outside the fovea to the edges of the retina there is a relatively consistent cone density, however, retinal ganglion cell density continues to decrease, leaving the retinal ganglion cells at the edge of the fovea with much larger receptive fields (Rosenholtz, 2016, Wässle et al., 1990). What this means is that the more eccentric parts of the visual field are more prone to crowding effects, leading to less discriminability for stimuli in the periphery (Toet and Levi, 1992). It has even been shown that the discriminability of summary statistic representations of visual search paradigms in proportion to those calculated by the peripheral retina directly predicts the efficiency of visual search (Rosenholtz et al., 2012). Our definition of attention as the functioning of a dynamic priority map argues that visual search efficiency is not an attentional constraint but rather a restrain of the acuity of the peripheral retina.

Where our definition of attention as the functioning of a dynamic priority map leaves us is that attention is not just a system meant only to reduce information, but rather a system meant to reduce information according to the current goals and intentions of the organism, which fits much more in line with our intuitions of attention as something that we can pay, hold, or lose. It also enables/forces us to review what we think we know about attention[32] in a way which I think could be very insightful and very productive.

Lastly, the definition of attention as a dynamic priority map provides a clear avenue for studying how semantic knowledge of objects, such as their size or their manipulability, can affect the functioning of attention. With the information of the dynamic priority map feeding back into lower-level visual circuits we can then begin to

[32] How attentional is visual search? What about inattentional blindness (see Rock et al., 1992)?

ask how that attentional feedback affects different circuits (perhaps across the two visual streams) in different ways by virtue of different connectivity patterns or competition with other inputs to the lower-level circuits. To put a metaphor on this, in the grand tradition of attentional filters or spotlights, attention is a hand. It may work differently when it is picking up a hammer versus when it is trying to pick up an elephant.

Attention and Platypuses

Shomstein, S., Zhang, X.L, and Dubbelde, D. (2022)

For this issue of WIREs we were invited to provide a perspective on "what is attention?" As researchers who study attention, our papers usually follow a similar script: give an example of a sensory overload, then appeal to the need to have a mechanism in place to sub-select some information from the overload for further processing, then define attention as selection. For example, in the most recent review from our lab (Shomstein et al., 2019) we start by posing the question of how does the human visual system sort through the massive amounts of sensory input, which it samples almost continuously, to arrive at a coherent perception of a scene? Then, we make the point that this process of searching through the environment for information is a ubiquitous component of sensory processing and it reflects a remarkable ability of the perceptual system to dynamically select information that is relevant for the current goal of the organism. Then, we finally get to the definition that the act of perceptual selectivity is what we call attention, and that this selection process is central to cognition. We could stop right here, since we defined attention as a selection mechanism, but this would hardly be satisfying to the reader, nor frankly to us.

Defining attention as a selection process is likely not going to stir much controversy, especially since this stems from William James' remark that attention is the "… taking possession by the mind, in clear and vivid form, of one out what seem several simultaneously possible objects or trains of thought" (James & Drummond, 1890). However, where things start to break down is when we try to further deconstruct what

selection is. What is being selected? How is it selected? What does it mean for information to be selected? As a scientific field we emphasize specificity and thus place value on putting forth testable models, and attention is no exception. Models of attention are many and focus on different aspects of the selection mechanism (similar to selection questions mentioned above). Some models focus on where the bottleneck of selection is, for example, early versus late and perceptual load theories of attention (Benoni & Tsal, 2013; Broadbent, 1958; Lavie et al., 2004; Makovski et al., 2014; Mevorach et al., 2014). Others focus on factors that guide selection, for example, top-down and bottom-up sources, and history (Awh et al., 2012; H. E. Egeth & Yantis, 1997; Itti & Koch, 2000). Yet, others focus on what it means for the signals in early sensory regions to instantiate selection, for example, divisive normalization (Denison et al., 2021; Li et al., 2015; Ni & Maunsell, 2017). This list of models is in no way meant to be exhaustive, but offers a small flavor of the kinds of aspects of attentional selection that are at the forefront of mechanistic investigations. Models are extremely helpful in helping fine tune our understanding of attention, and have been proven to be extremely fruitful in terms of driving inquiry. One concern with most of these models, however, is the issue of falsifiability. Few of these models clearly outline conditions under which the models will be clearly falsifiable, thus, when a new discovery or phenomenon pops up that does not easily fit into this or that theory, there is a tendency, for some, to raise a righteous finger and say “see, you have ill-defined theories of attention, and thus you do not even know what you are studying!” This is where we think it is important to remind ourselves that: (i) what we are doing is reverse engineering a system, the workings of which are currently at the level of hypotheses; (ii) outlining models is necessary in order to arrive at

the ultimate ‘correct answer’ of how the system works; (iii) current models are just that – current – and are by definition flawed; and (iv) if a model is found to be insufficient, it is no reason to claim ignorance in what we are doing, rather, it is a reason to tweak the models and continue testing and improving them; and finally (v) at some point it might no longer be useful to hold on to a particular model as it might feel like trying to fit a square peg into a round hole. It is this last point that the title of our opinion piece is making a reference to—the attentional platypus.

WHAT'S WITH THE PLATYPUS?

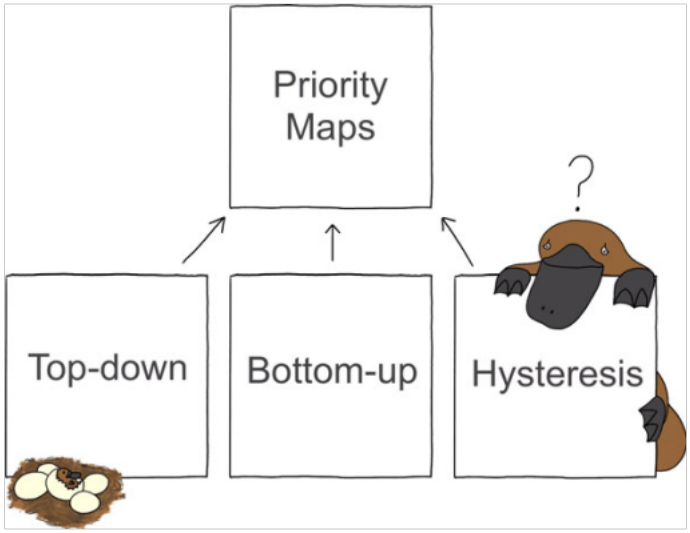
As an example, let us consider the top-down and bottom-up model of attention. Until relatively recently, reviewing decades of behavioral research led to a synthesis that the distribution of attention is largely guided by intentions of the observer (top-down) as well as by the salience of the physical stimulus (capture or bottom-up). A classic example of capture is a single raspberry in a bowl of blueberries, or the siren of a fire truck. In both cases, salience is defined by signals that are sufficiently different from the sensory environment they are embedded in by standing out from the context. On the flip side of capture, or bottom-up attention, is attentional guidance that is under full control of the organism. An example of such top-down attentional control is focusing on a particular aspect of a sensory environment. For example, when meeting a friend in a crowded square (before COVID), knowing that your friend is wearing a red t-shirt will guide your search to focus predominantly through people who are wearing red (H. E. Egeth & Yantis, 1997; J. M. Wolfe, 2021). This is an example of feature-based selection (Liu, 2019; Treisman & Gelade, 1980). Another type of top-down guidance is space-based

selection: knowing that your friend is meeting you at the NW corner, you will directly attend that specific spatial location on the crowded square (Posner, 1980). More recently, a large body of attention phenomena have emerged that could not be easily accommodated into a top-down and bottom-up guidance network (Awh et al., 2012). These phenomena include things like guidance of attention based on various expectations either set by prior associated value, prior knowledge of spatial occurrence, or by various contingencies of the stimulus (Anderson et al., 2011; Drummond & Shomstein, 2010; Geng & Behrmann, 2002, 2005; Moore & Egeth, 1998; Shaw, 1978; Shomstein & Yantis, 2004). In response to the amassed volume of attentional platypuses (evidence), the traditional two-factor top-down and bottom-up attentional guidance theory was revised to accommodate a third factor—hysteresis. This revision was indeed a welcomed change as now a whole host of attentional phenomena have found their home in an attentional model, thus paving the way for new investigations of this tri-factor model of attentional control (Awh et al., 2012). This new approach ushered in a renewed interest into attentional selection, and the paper itself resonated with researchers, having been cited over 1300 times. Several follow-up review articles that further clarified the tri-factor model followed (e.g., see Theeuwes, 2018a) and a set of commentary papers that followed (Becker, 2018; Chelazzi & Santandrea, 2018; H. Egeth, 2018; Gaspelin & Luck, 2018; Kryklywy & Todd, 2018; Sisk et al., 2018; Theeuwes, 2018b; J. Wolfe, 2018). We argue, however, that while this approach has been fruitful, it is still limited and it compresses disparate effects together for the sake of organizational clarity. This is akin to evolutionary taxonomy, and it is where taking a cue from the animal kingdom might be helpful. Traditionally, among the vertebrates, there are fish, amphibians, reptiles, birds,

and of course the mammals. This is a compelling organizational structure, it explains most of the variance among the vertebrates that you will ever see, and for this reason it is still taught in elementary schools. Digging a bit deeper though, it begins to fall apart. Where are we to put platypuses with the fur of a mammal but the eggs and beak of a bird? Where do we put lungfishes with the body of a fish but the lungs of an amphibian? Quite a few attentional phenomena can be called attentional platypuses in their inability to fit neatly into the tri-factor model.

Recent research has shown several examples of potential attentional platypuses, attentional effects that do not seem to fit into any of the three traditional boxes of the tri-factor model: top-down, bottom-up, and hysteresis. These effects include most attentional phenomena associated with task-irrelevant sensory information impinging on attentional control (e.g., semantic relatedness, object size, spontaneous cognitive states, rhythmic fluctuations, strategies). These platypus effects are neither top-down or bottom-up, and they are not easily explained by hysteresis (without stretching the concept to an almost uncomfortable all-inclusive catchall third box), but they nevertheless exert influence on attentional orienting. For example, the influence of task-irrelevant items on attentional orienting could be argued to be top down, as the target location can be defined by its spatial relationships to the nontargets, but this could also be argued as hysteresis, as these relationships are only known after seeing the spatial structure of the trial. If task-irrelevant influences on attention are mammals, and hysteresis effects are birds, then this effect of non-targets is certainly a platypus. In the next few sections, we will exemplify a few instances of attentional effects that fit poorly (attentional platypuses, see Figure 1) into the tri-factor model of attention, and then suggest a possible way forward, the

dynamically weighted prioritization approach, that is more flexible in terms of accommodating attentional platypuses. While we do not argue against the tri-factor model, and appreciate the benefit of organizing attentional selection in these terms, we submit that a more detailed model of attentional taxonomy could liberate us, as a field, to yield new investigations and lead to new models.



EXAMPLES OF SOME PLATYPUSES

Task-irrelevant object size

Size is an intrinsic attribute of all objects in the physical world, and its computation is inherently present given the retinotopic nature of visual processing in the ventral visual system (i.e., size of the retinal image and observer's distance) (Baird, 1963; Hubbard et al., 1989). Castiello and Umiltà (1990) demonstrated that if participants were asked to respond to a presence of a target, their response times were modulated by the size of the box in which the target was embedded. In a more recent experiment, Nah et al. (2018) demonstrated that size can also affect attentional shifting. Using an adapted Egly paradigm (Egly et al., 1994) in which the classic rectangles were replaced by trapezoids,

attentional shifting was slower when shifting to the thin end of a trapezoid than when shifting to the thick end. The authors argue that this is a manifestation of Fitt's law, the phenomenon known from the motor system in which larger objects are more quickly acted upon than smaller objects (Fitts, 1954). Both studies demonstrate that size is an inherent property of an object and that it has ramifications on the functioning of attention. In either of these cases, size of the object was irrelevant to the task at hand, neither was it a salient feature, nor was there a particular history associated with any of the objects that is unique under large or small instances. As such, task-irrelevant size of the object, and its influence on attentional allocation, is not easily accommodated by current theories of attentional control.

Furthermore, Collegio et al. (2019) demonstrated that semantically known size, rather than retinal size, impacts attentional allocation. In their paradigm, a line drawing of a known object, such as a domino or a kayak, was presented at a fixed retinal size at fixation. One end of the object was then cued and a target appeared either at the cued end of the object, or on the other side of the object. It was found that participants were slower to respond to the invalid target when the object was known to be large (e.g., kayak) than when it was known to be small (e.g., domino). The influence of the inferred object size was observed despite the exact same retinal size in which all objects were presented. In both instances of object size influencing attentional allocation (physical size differences and object semantic size), size was a task-irrelevant factor that influenced attentional allocation and is a factor that is not easily accommodated by a tri-factor model.

Task-irrelevant semantic relationships

In recent years there has been a renewed interest in findings showing that semantic and category information is rapidly processed (Biederman et al., 1974; Greene & Oliva, 2009; Potter, 1975) in a manner that minimally engages attentional selection. Several studies extended this observation to suggest that semantic and category information of items maintained in working memory tend to influence attentional allocation. For instance, a search target (e.g., a chimney) stored in visual working memory (VWM) facilitates the directing of attention toward likely positions in a scene (a roof) with the very first eye movement (Castelhano & Heaven, 2010; Eckstein et al., 2006; Malcolm & Henderson, 2010; Neider & Zelinsky, 2006). Similarly, an object stored in VWM (e.g., a motorcycle) increases the likelihood that a viewer will attend to semantically related objects (a motorcycle helmet) (Belke et al., 2008; Hwang et al., 2011; Mack & Eckstein, 2011; Moores et al., 2003), or that semantically related distractors will capture attention (Belke et al., 2008; Moores et al., 2003).

In the cases described above where task-relevant semantic and category information guides attentional selection, these guiding factors can be easily placed into the hysteresis box of the tri-factor model of attention. However, the attentional platypus appears when semantic/category factors become irrelevant to the task as it no longer fits into top-down or hysteresis, and it is not capture. In a recent study, Malcolm et al., 2016, showed that when presented with task-irrelevant objects, the object semantic relatedness influences attention even if object identity and semantic relationships are not predictive of the task at hand. In a set of follow-up functional magnetic resonance imaging (fMRI) experiments, Nah et al. (2021) fleshed out the neural network that is responsible for task-irrelevant semantic influence on attentional selection. It was shown that the left inferior

frontal gyrus (IFG) shows sensitivity to object semantic relatedness, with activity in IFG directly predicting the degree of behavioral benefit of faster response times for targets that appear on task-irrelevant semantically related objects. It was also shown that semantic relatedness then likely biases spatial attention maps in the intraparietal sulcus, subsequently modulating early visual cortex activity. These results suggested that if an object is attended, its semantic properties bias attention, even if it is irrelevant to the ongoing task or if more predictive factors are available, outlining a factor of attentional selection not easily accommodated by prior models of attentional selection. Taken together, our results, and those of others, show that the semantic content and the relative semantic relationships, even while task irrelevant, exert their influence on attentional allocation and is a factor that is not easily accommodated by a tri-factor model.

Spontaneous cognitive states

Along the lines of attention research focusing on external factors (such as physical salience, task goals from explicit instruction, the in-flowing experience during the experimental session, as well as task-irrelevant size and semantics discussed in previous sections), there have been other lines of research looking at the degree to which internal factors, such as spontaneous cognitive states, influence attention. The latter cannot be intuitively categorized as bottom-up, top-down, or history-driven processing. Although these spontaneous factors have not been included in the tri-factor model of attention, there is considerable evidence in the literature that these seemingly task-irrelevant factors nonetheless play a role during attention deployment. Several recent theories argue that cognitive states influence attention and perception (Goldstone et al., 2015; Lupyan, 2017;

although there is ongoing debate about whether cognitive states actually influence judgments or responses instead of perception per se, Firestone & Scholl, 2016). For example, knowledge can greatly influence the perception of bistable or ambiguous images (Meng & Tong, 2004; Tong et al., 1998). Attention is also known to be biased toward emotional materials, interacting with the emotional states of the individuals (see a review in Yiend, 2010). For example, patients with generalized anxiety disorder (GAD) experienced more interference on a cognitive task from threat words occurring in the unattended auditory channel, compared to control participants (Mathews & MacLeod, 1986). This finding suggests that patients with GAD may prioritize threat words differently from control participants. While emotional stimuli can be attributed to hysteresis defined as a broader term (e.g., "previous experience" instead of narrowly "past selection history" as defined in Awh et al., 2012), the emotional state of the observer cannot easily be accommodated. Similarly, intention or voluntary action is another internal factor that can guide our attention allocation, separate from top-down attention modulated by task goals. For example, participants could voluntarily shift attention between two locations and these voluntary, not-task restricted shifts of attentional focus have been successfully tracked by neural activities in the brain (Gmeindl et al., 2016). Furthermore, attention to intention has been found to have neural substrates in prefrontal and presupplementary motor areas (Lau et al., 2004), indicating the biological possibility for the influence of internal intention on attention. These spontaneous processes of knowledge, emotions, intention, and others are typically not explicitly instructed and can be considered as task-irrelevant within current attentional

models. Their influences on attentional selection are supported by growing evidence, but may have been inadvertently left out of the more traditional definitions of attention.

Rhythmic fluctuations

There is an important assumption that is often made in many attention studies, as well as in the tri-factor model: that the attentional states are stable throughout the course of the experiment session. This simplification on the one hand helps with noise reduction and theoretical interpretation of the results, but on the other hand hides the dynamic nature of the attentional processes. If we look specifically at the temporal dimension, we can find rhythmic oscillations in attentional performance. For example, with explicit instruction to sustain attention at one location, participants show lapses in their responses in rhythmic patterns (Smith et al., 2003), which may indicate fluctuations of attentional prioritization. These behavioral patterns may be tied with intrinsic neural rhythms in the brain, especially those in attention-related regions (Fiebelkorn & Kastner, 2019) and activities in the default-mode network (Sali et al., 2016). This spontaneous fluctuation of attention is sometimes explained by one common and long-studied phenomenon called inhibition of return (IOR)—the tendency that spatial attention is inhibited to orient to the previously attended location (Posner et al., 1985). However, the studies of the IOR phenomenon itself do not often discuss it in the context of temporal attention on an elongated scale, and the temporal characteristics are rarely highlighted while discussing the definitions and models of attention. We propose that comprehensive understanding would require further exploration of these aspects as well.

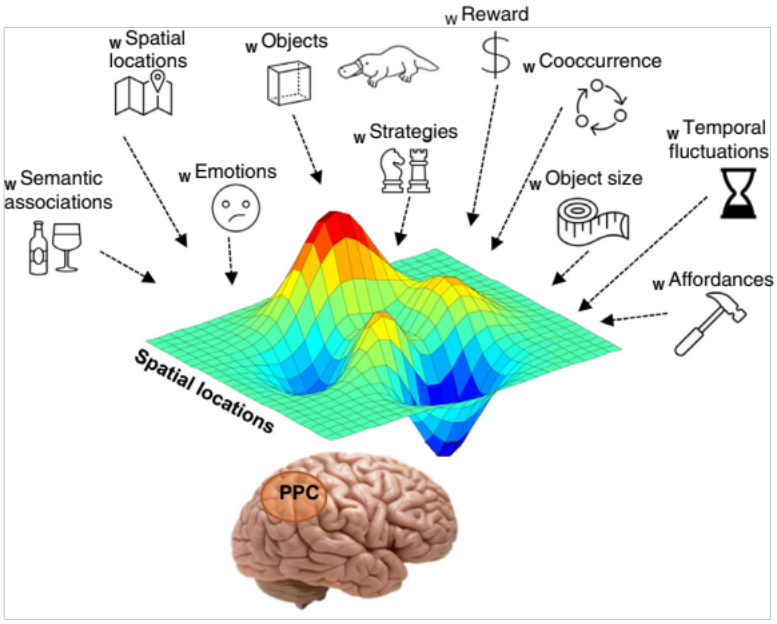
Strategies

Strategy is another higher-level cognition that has been proposed as playing a role in attentional selection. With the same instruction, participants may come up with different strategies to accomplish the task, which differs where and how their attention is allocated. For example, in a visual search paradigm where the optimal strategy is to change the color to search from trial to trial, participants tend to be sluggish to update the optimal color and there are substantial individual differences in using this optimal strategy (Irons & Leber, 2016). In some circumstances, participants' strategies may not even align with the task requirements. It has been found that attention sometimes may be geared toward objects or space which is neither beneficial to the current task goal nor of immediate perceptual priority, for example, toward novel stimuli (Becker & Horstmann, 2011; Ernst et al., 2020; Retell et al., 2016). Some researchers propose this behavior pattern as a spontaneous act for novelty seeking, which is important for exploration (Kakade & Dayan, 2002). A relevant behavioral pattern comes from the load theory of attention: when perceptual load is low, participants pay more attention to the distractors, compared to the high perceptual load condition where perceptual capacity is exhausted on task-relevant stimuli (Lavie et al., 2004; a more recent review in Murphy et al., 2016). This may indicate a hardwired mechanism that makes it difficult to avoid attending to task-irrelevant stimuli in order to fill up perceptual capacity. Higher-level strategy or hardwired cognition is another factor that cannot be easily categorized in the tri-factor model of attention and our understanding of these mechanisms benefits little from simply categorizing them without studying their interactions with other attentional factors.

WHAT DO WE LEARN FROM THE PLATYPUSES?

Studying attention is not too different from studying evolution. The way we learn about the evolution of life is not just by studying a single model organism in isolation, but rather by taking a wider view and observing the commonalities in the development of many different organisms. Similarly, while it was extremely useful to focus on developing general attentional models, an understanding of attention as a larger process requires sensitivity to the full range of factors (including the platypuses) that ultimately influence attentional allocation. The focus should now shift to not only delineating which factors influence attention, but examining similarities and differences between them, and, perhaps most importantly, on understanding how they interact.

Attentional research has amassed an incredible amount of knowledge regarding the general ways in which attention is allocated as well as the various factors that influence attention. We propose that the time is ripe to consider attentional allocation as a state in the dynamic priority map (Figure 2) which takes as input various factors that are known to influence attentional allocation (Bisley & Goldberg, 2010; Koch & Ullman, 1987; also see reviews Ptak, 2012; Shomstein & Gottlieb, 2016; Todd & Manaligod, 2018). The state dynamically changes as a function of multiple attentional factors that influence it, with corresponding weights that determine the degree of influence for that particular factor. This approach enables researchers to launch inquiries about the interactions between and among factors that influence attentional allocation. An added benefit of this approach is that no sensory stimulus is thought of in binary terms as task-irrelevant or not salient, rather, all sensory stimuli are contributing to attention and thus lack the need for a particular binary label or belonging status.



The current state of the priority map is likened to a Bayesian framework (Feldman & Friston, 2010; Friston, 2009) in which prior and current information about a particular attentional factor's reliability modulates the weight of that factor as a constraint on the attentional priority map in posterior parietal cortex (Figure 2). Representations with a high weighting successfully contribute more to constraining attention, and factors with less weight contribute less. The weighting changes dynamically reflecting the current state of the system. As the weightings of some representations are reduced due to an increase in spatial uncertainty (e.g., you are less sure about where the target is), for example, other representations (e.g., objects) contribute to a greater extent, and therefore constrain attentional selection (Donovan et al., 2017; Drummond & Shomstein, 2013). We call this the Dynamic Prioritization Approach to predicting attentional allocation. Of course, we are perfectly aware that what we are arguing for here is ultimately a proposition to "study everything at once" and understand the challenge that is associated

withit. However, more and more empirical evidence has accumulated showing that there is no such thing as an irrelevant sensory stimulus (see Shomstein et al., 2019 for most recent review). Perhaps one practical step in the process is to continue investigating the relative contribution of two attentional factors at once and their interactions (e.g., spatial locations and objects; objects and reward; certainty and emotion) and move on to multiple-way interactions. At the same time, investigating the profile of a complex and dynamic system from a zoom-out perspective can provide additional helpful insights.

Moreover, this dynamically weighted prioritization approach lends itself nicely to a somewhat neglected investigation of individual differences in attentional orienting. In a similar set of circumstances, attentional allocation (e.g., the state of the priority map) in two individuals might be different because of different weights associated with some or all of the attentional factors. This approach allows for not only studying what is similar among individuals (i.e., what factors are involved) but also how the relative contribution of each factor is determined through differential weighting according to individual differences. We should note that the dynamically weighted prioritization approach that we put forward here suffers from a similar problem as other models of attention do—it is unclear whether it presents a clear case for falsifiability. Currently, our lab is on a mission to create a set of hypotheses that follow directly from this approach to test several boundary conditions that will ultimately help with figuring out conditions of falsifiability.

CONCLUSION

In a way, the platypus metaphor is not a metaphor at all. The systems we are studying evolved in the exact same manner as platypuses (Cisek, 2019). If platypuses do not fall into neat boxes with birds or mammals, analogously there is no reason why factors that influence attention (those listed in Figure 2, for example) should either. When attentional platypuses are encountered in the "wild" we should not be trying to force them into groups with the birds or the mammals, but rather re-evaluate taxonomies. Going forward, we suggest that we think more in terms of continuities and dynamics. In essence, we should not force an arbitrary organization on the menagerie of attentional mechanisms which we have, but rather we should use the organization endemic to the system itself, prioritization.

CHAPTER 5 – Conclusion

The theoretical work and the experiments presented throughout this dissertation demonstrate several ways in which what we know about an object change how we perceive that object through altering the nature of the object's representation in neural circuits. The first paper presented, "Object Perception Is Modulated By Size, Both Actual And Semantic", supports the hypothesis that the semantic size of an object is encoded into the object's representation by presenting data showing that an effect of presentation size (less efficient processing for targets in larger objects) can also be elicited by semantic size. Not only does this evidence semantic knowledge altering perception, but it also suggests a possible mechanism by which semantic effects can arise. Namely, behavioral effects which are elicited by low-level features, such as size, become associated to the semantic knowledge of the objects composed of such low-level features. The second paper, "Mugs And Plants: Object Semantic Knowledge Alters Perceptual Processing With Behavioral Ramifications", demonstrates that the neural circuits that are engaged in the perception of an object, which depends on the semantic identity of the object, has repercussions for how the object is perceived. Manipulable objects, which elicit neural activity in the dorsal visual processing stream, are perceived with higher temporal resolution, while non-manipulable objects, more heavily reliant on the temporal visual processing stream, are perceived with higher spatial resolution. Lastly, "Attention and Platypuses" presents an argument for how to think about attention, defining attention specifically as the functioning of a dynamic priority map. This provides a better framework for thinking about the myriad of attentional effects than the traditional box plot categorization schemes. Additionally, it suggests two ways in which semantic

knowledge about objects can impact the functioning of attention. One, semantic knowledge can directly impact the functioning of the dynamic priority map. Two, semantic knowledge can alter where across the cortex an object's representation is processed which has implications for how effective the attentional process is at altering activity in those circuits.

The effect of object semantics on perception and attention is a major focus of study both in the work in the Attention and Cognition Laboratory but also across Cognitive Psychology. However, if we bring the concepts of object semantics from Cognitive Psychology into Cognitive Neuroscience by examining the underlying neural mechanisms we can more deeply understand what it means for us to know about an object. For example, in Chapter 3 I present evidence that the known manipulability of an object alters how we perceive the object because manipulability toggles the engagement of the two visual streams dependent on their evolved purposes. Manipulability is not just a semantic category in the case, but also a key factor driving the evolutionary development of the two visual streams[33], and so the repercussions of manipulability on our perception and therefore behavior is deeply rooted in the ontogeny of our brains. This sort of evolutionary perspective on object semantics will not be able to bring us to a complete understanding of what object semantics are but it can help us to understand a number of important aspects of our psychology, such as:

- What about how we experience semantic knowledge is innate?
- What about our semantic experience is uniquely human and do we share any of the associated mechanisms with other creatures?

[33] At least in primates.

- Can an understanding of the evolutionary pressures which drove the development of systems like our visual system be generalized to broadly applicable principles of biological information processing systems such that we can better understand non-human intelligence[34].

Before I stray too far, I set out to argue that object representations are not all the same, that they differ depending on things like semantic size or known manipulability such that our perception of the object is altered. If I have not convinced you that object representations do differ, I hope that I have at least convinced you that it is an interesting possibility. One last parting thought. With me being both (1) myself and also (2) an attention researcher, it is absolutely necessary to put a metaphor on top of the central thesis of this dissertation. The same neural circuits involved in lower-level visual perception are activated in both the perception of a visual stimulus and in the holding of that visual stimulus in working memory (Teng and Kravitz, 2019; Adam et al., 2022). In other words, the representation of an object is accessed both by perception of the object and by thinking of the object, meaning that our perception of an object is changed depending on how what we know about the object changes its representation. Think of Plato's cave[35] in which people are chained underground in such a way that they can only see the shadows cast onto the wall before them. The shadows are their entire reality. Similar to how the people in the cave never saw the objects casting the shadows, we never hold any actual objects in our heads[36]. What we have is shadows of the objects

[34] I am 100% talking about cetaceans and cephalopods and not extraterrestrials. Don't be weird.
[35] Activate your 'Plato's cave' representation.
[36] Unless you're Phineas Gage.

captured in the activity of our neural circuits. Similar to how the irregular walls of the cave will distort the shadows, the physiology of our neural circuits distorts our representations. Depending on where they are on the cave wall, or where they are in our brains, the shadows are not all the same.

REFERENCES

Abbie, A. A. (1940). Cortical lamination in the monotremata. *Journal of Comparative Neurology*, *72*(3), 429-467. https://doi.org/10.1002/cne.900720302

Abbie, A. A. (1942). Cortical lamination in a polyprotodont marsupial, Perameles nasuta. *Journal of Comparative Neurology*, *76*(3), 509-536. https://doi.org/10.1002/cne.900760310

Abrams, R.A. and Weidler, B.J. (2014). Trade-offs in visual processing for stimuli near the hands. *Attention, Perception, and Psychophysics*, 76, 383-390. https://doi.org/10.3758/s13414-013-0583-1

Adam, K., Rademaker, R. L., & Serences, J. (2021). Evidence for, and challenges to, sensory recruitment models of visual working memory. https://doi.org/10.31234/osf.io/wb5e6

Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(39), 15214–15218. https://doi.org/10.1073/pnas.0805867105

Almeida, J., Mahon, B. Z., & Caramazza, A. (2010). The role of the dorsal visual processing stream in tool identification. *Psychological Science: A Journal of the American Psychological Society*, *21*(6), 772–778. https://doi.org/10.1177/0956797610371343

Almeida, J., Fintzi, A. R., & Mahon, B. Z. (2013). Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. *Cortex*, *49*(9), 2334–2344. https://doi.org/10.1016/j.cortex.2013.05.004

Almeida, J., Mahon, B.Z., Zapater-Raberov, V., Dziuba, A., Cabaço, T., Marques, F., and Caramazza, A. (2014). Grasping with the eyes: The role of elongation in visual recognition of manipulable objects. *Cognitive, Affective, & Behavioral Neuroscience*, 14, 319-335. https://doi.org/10.3758/s13415-013-0208-0

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25), 10367–10371. https://doi.org/10.1073/pnas.1104047108

Andersen, P., & Lømo, T. (1966). Mode of activation of hippocampal pyramidal cells by excitatory synapses on dendrites. *Experimental Brain Research*, *2*(3), 247-260. https://doi.org/10.1007/BF00236716

Anderson, S. R., & Wiens, J. J. (2017). Out of the dark: 350 million years of conservatism and evolution in diel activity patterns in vertebrates. *Evolution*, *71*(8), 1944-1959. https://doi.org/10.1111/evo.13284

Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *Journal of neuroscience*, *29*(34), 10638-10652. https://doi.org/10.1523/JNEUROSCI.2807-09.2009

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Baird, J. C. (1963). Retinal and assumed size cues as determinants of size and distance perception. *Journal of Experimental Psychology*, 66(2), 155–162. https://doi.org/10.1037/h0046554

Baizer, J. S., Ungerleider, L. G., & Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, *11*(1), 168–190. https://doi.org/10.1523/jneurosci.11-01-00168.1991

Barnas, A.J. & Greenberg, A.S. (2016). Visual field meridians modulate the reallocation of object-based attention. *Atten Percept Psychophys* 78:1985–1997 (2016). https://doi.org/10.3758/s13414-016-1116-5

Barnas, A.J. & Greenberg, A.S. (2019). Object-based attention shifts are driven by target location, not object placement. *Visual Cognition*, 27:768-791, https://doi.org/10.1080/13506285.2019.1680587

Becker, S. I. (2018). Reply to Theeuwes: Fast feature-based top-down effects, but saliency may be slow. *Journal of Cognition*, 1(1), 28. https://doi.org/10.5334/joc.23

Becker, S. I., & Horstmann, G. (2011). Novelty and saliency in attentional capture by unannounced motion singletons. *Acta Psychologica*, 136(3), 290–299. https://doi.org/10.1016/j.actpsy.2010.12.002

Belke, E., Humphreys, G. W., Watson, D. G., Meyer, A. S., & Telling, A. L. (2008). Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, 70(8), 1444–1458. https://doi.org/10.3758/PP.70.8.1444

Benoni, H., & Tsal, Y. (2013). Conceptual and methodological concerns in the theory of perceptual load. *Frontiers in Psychology*, 4, 522. https://doi.org/10.3389/fpsyg.2013.00522

Besner, D. (1983). Visual pattern recognition: Size preprocessing re-examined. The Quarterly Journal of Experimental Psychology, 35(1), 209–216. https://doi.org/10.1080/14640748308402126

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103(3), 597–600. https://doi.org/10.1037/h0037158

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, 33, 1–21. https://doi.org/10.1146/annurev-neuro-060909-152,823

Blasdel, G. G., & Lund, J. S. (1983). Termination of Afferent Axons in Macaque Striate Cortex. *Journal of Neuroscience*, *3*(7), 1389–1413. https://doi.org/10.1523/JNEUROSCI.03-07-01389.1983

Bocanegra, B.R. and Zeelenberg, R. (2009). Emotion Improves and Impairs Early Vision. *Psychological Science*, 20(6), 707-713. https://doi.org/10.1111/j.1467-9280.2009.02354.x

Broadbent, D. E. (1958). Perception and communication. *Pergamon Press*. http://search.ebscohost.com/direct.asp?db=pzh&jid=%22200416224%22&scope=site

Brown, J. F. (1931). The visual perception of velocity. Psychologische Forschung, 14(1), 199–232. https://doi.org/10.1007/BF00403873

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. Journal of Experimental Psychology: Human Perception and Performance, 1(3), 214–220. https://doi.org/10.1037/0096-1523.1.3.214

Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics*, 72(5), 1283–1297. https://doi.org/10.3758/APP.72.5.1283

Castiello, U. and Umilta, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica* 73(3), 195-209. https://doi.org/10.1016/0001-6918(90)90022-8

Chan, A.W.Y., Kravitz, D.J., Truong, S., Arizpe, J., and Baker, C.I. (2010). Cortical representations of bodies and faces are strongest in commonly experienced configurations. *Nature Neuroscience*, 13(4): 417-418. https://doi.org/10.1038/nn.2502

Chan, D., Peterson, M., Barense, M., and Pratt, J. (2013). How action influences object perception. *Frontiers in Psychology*, 4. https://doi.org/10.3389/fpsyg.2013.00462

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

Chelazzi, L., & Santandrea, E. (2018). The time constant of attentional control: Short, medium and long (infinite?). *Journal of Cognition*, 1(1), 27. https://doi.org/10.5334/joc.24

Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Attention, Perception, & Psychophysics*, 81(7), 2265–2287. https://doi.org/10.3758/s13414-019-01760-1

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal

characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291-307. https://doi.org/10.1093/brain/123.2.291

Collegio, A.J., Nah, J.C., Scotti, P.S., & Shomstein, S. (2019). Attention scales according to inferred real-world object size. *Nature: Human Behavior,* 3:40–47. https://doi.org/10.1038/s41562-018-0485-2

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Connolly, M., & Van Essen, D. (1984). The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology*, *226*(4), 544-564. https://doi.org/10.1002/cne.902260408

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature neuroscience*, *3*(3), 292-297. https://doi.org/10.1038/73009

Dart, R. A. (1934). The dual structure of the neopallium: Its history and significance. *Journal of Anatomy*, *69*(Pt 1), 3.

De Haas, B., Schwarzkopf, D.S., Alvarez, I., Lawson, R.P., Henriksson, L., Kriegeskorte, N., and Rees, G. (2016). Perception and Processing of Faces in the Human Brain Is Tuned to Typical Feature Locations. *The Journal of Neuroscience*, 36(36):9289-9302. https://doi.org/10.1523/JNEUROSCI.4131-14.2016

DeLoache, J. S. (2004). Scale Errors Offer Evidence for a Perception-Action Dissociation Early in Life. Science, 304(5673), 1027–1029. https://doi.org/10.1126/science.1093567

Denison, R. N., Carrasco, M., & Heeger, D. J. (2021). A dynamic normalization model of temporal attention. *Nature Human Behaviour*, 1–12, 1674–1685. https://doi.org/10.1038/s41562-021-01129-1

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*(1), 219–240. https://doi.org/10.1113/jphysiol.1984.sp015498

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J. O. N., Cox, R., ... & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences*, *93*(6), 2382-2386. https://doi.org/10.1073/pnas.93.6.2382

DeYoe, E. A., & van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*(5), 219–226. https://doi.org/10.1016/0166-2236(88)90130-0

Dhande, O. S., Stafford, B. K., Lim, J. H. A., & Huberman, A. D. (2015). Contributions of retinal ganglion cells to subcortical visual processing and behaviors. *Annual review of vision science*, *1*, 291-328. https://doi.org/10.1146/annurev-vision-082114-035502

Diamond, R., & Carey, S. (1986). Why Faces Are and Are Not Special. An Effect of Expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Donovan, I., Pratt, J., & Shomstein, S. (2017). Spatial attention is necessary for object-based attention: Evidence from temporal-order judgments. *Attention, Perception, & Psychophysics*, 79(3), 753–764. https://doi.org/10.3758/s13414-016-1265-6

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470-2473. https://doi.org/10.1126/science.1063414

Drummond, L., & Shomstein, S. (2010). Object-based attention: Shifting or uncertainty? *Attention, Perception, & Psychophysics*, 72(7), 1743–1755. https://doi.org/10.3758/APP.72.7.1743

Drummond, L., & Shomstein, S. (2013). The timecourse of space- and object-based attentional prioritization with varying degrees of certainty. *Frontiers in Integrative Neuroscience*, 7, 88. https://doi.org/10.3389/fnint.2013.00088

Dubbelde, D. and Shomstein, S. (2022). Mugs and Plants: Object Semantic Knowledge Alters Perceptual Processing with Behavioral Ramifications. *Psychological Science, accepted.*

Eckstein, M. P., Koehler, K., Welbourne, L. E., & Akbas, E. (2017). Humans, but Not Deep Neural Networks, Often Miss Giant Targets in Scenes. *Current Biology*, 27(18), 2827-2832.e3. https://doi.org/10.1016/j.cub.2017.07.068

Egeth, H. (2018). Comment on Theeuwes's characterization of visual selection. *Journal of Cognition*, 1(1), 26. https://doi.org/10.5334/joc.29

Egeth, H. E., & Yantis, S. (1997). Visual Attention: Control, representation, and time course. *Annual Review of Psychology*, 48(1), 269–297. https://doi.org/10.1146/annurev.psych.48.1.269

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*(2), 161–177. https://doi.org/10.1037/0096-3445.123.2.161

Emmert E. (1881) Größenverhältnisse der Nachbilder. Klinische Monatsblätter für Augenheilkunde und für augenärztliche Fortbildung, 19: 443-450.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598-601. https://doi.org/10.1038/33402

Ernst, D., Becker, S., & Horstmann, G. (2020). Novelty competes with saliency for attention. *Vision Research*, 168, 42–52. https://doi.org/10.1016/j.visres.2020.01.004

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8*(10), 1380–1385. https://doi.org/10.1038/nn1537

Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-Dependent Representation of a Size Illusion in Human V1. Current Biology, 18(21), 1707–1712. https://doi.org/10.1016/j.cub.2008.09.025

Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 215. https://doi.org/10.3389/fnhum.2010.00215

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, *358*(6389), 756–758. https://doi.org/10.1038/358756a0

Fiebelkorn, I. C., & Kastner, S. (2019). A rhythmic theory of attention. *Trends in Cognitive Sciences*, 23(2), 87–101. https://doi.org/10.1016/j.tics.2018.11.009

Firestone, C., & Scholl, B. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences, 39*, E229. https://doi.org/10.1017/S0140525X15000965

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391. https://doi.org/10.1037/h0055392

Flandrin, J. M., & Jeannerod, M. (1981). Effects of unilateral superior colliculus ablation on oculomotor and vestibulo-ocular responses in the cat. *Experimental brain research*, *42*(1), 73-80. https://doi.org/10.1007/BF00235731

Freud, E., Plaut, D. C., & Behrmann, M. (2016). ‘What’ Is Happening in the Dorsal Visual Pathway. *Trends in Cognitive Sciences*, *20*(10), 773–784. https://doi.org/10.1016/j.tics.2016.08.003

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences*, 22(1), 79–92. https://doi.org/10.1016/j.tics.2017.11.001

Geng, J. J., & Behrmann, M. (2002). Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, 13(6), 520–525. https://doi.org/10.1111/1467-9280.00491

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, 67(7), 1252–1268. https://doi.org/10.3758/BF03193557

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. Journal of Neuroscience, 9(7), 2432–2422. https://doi.org/10.1523/jneurosci.09-07-02432.1989

Gmeindl, L., Chiu, Y.-C., Esterman, M. S., Greenberg, A. S., Courtney, S. M., & Yantis, S. (2016). Tracking the will to attend: Cortical activity indexes self-generated, voluntary shifts of attention. *Attention, Perception, & Psychophysics*, 78(7), 2176–2184. https://doi.org/10.3758/s13414-016-1159-7

Goldstone, R. L., de Leeuw, J. R., & Landy, D. H. (2015). Fitting perception in and to cognition. *Cognition*, 135, 24–29. https://doi.org/10.1016/j.cognition.2014.11.027

Gomez, M. A., Skiba, R. M., & Snow, J. C. (2018). Graspable Objects Grab Attention More Than Images Do. *Psychological Science*, 29(2), 206–218. https://doi.org/10.1177/0956797617730599

Goodale, M. A., Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, *4*(7), 604-610. https://doi.org/10.1016/S0960-9822(00)00132-9

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25. https://doi.org/10.1016/0166-2236(92)90344-8

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154-156. https://doi.org/10.1038/349154a0

Goulas, A., Margulies, D. S., Bezgin, G., & Hilgetag, C. C. (2019). The architecture of mammalian cortical connectomes in light of the theory of the dual origin of the cerebral cortex. *Cortex*, *118*, 244-261. https://doi.org/10.1016/j.cortex.2019.03.002

Gozli, D. G., West, G. L., & Pratt, J. (2012). Hand position alters vision by biasing processing through different visual pathways. *Cognition*, *124*(2), 244–250. https://doi.org/10.1016/j.cognition.2012.04.008

Graziano, M. S. A., & Gross, C. G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental Brain Research*, *97*(1), 96–109. https://doi.org/10.1007/BF00228820

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, 20(4), 464–472. https://doi.org/10.1111/j.1467-9280.2009.02316.x

Hall, M. I., Kamilar, J. M., & Kirk, E. C. (2012). Eye shape and the nocturnal bottleneck of mammals. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4962-4968. https://doi.org/10.1098/rspb.2012.2258

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory.* Wiley.

Heesy, C. P., & Hall, M. I. (2010). The nocturnal bottleneck and the evolution of mammalian vision. *Brain, behavior and evolution*, *75*(3), 195-203. https://doi.org/10.1159/000314278

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J. H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, *81*(7), 2288-2303. https://doi.org/10.3758/s13414-019-01846-w

Hubbard, T. L., & Baird, J. C. (1988). Overflow, First-Sight, and Vanishing Point Distances in Visual Imagery. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(4), 641–649. https://doi.org/10.1037/0278-7393.14.4.641

Hubbard, T. L., Kall, D., & Baird, J. C. (1989). Imagery, memory, and size-distance invariance. *Memory & Cognition*, 17(1), 87–94. https://doi.org/10.3758/BF03199560

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, *148*(3), 574. https://doi.org/10.1113/jphysiol.1959.sp006308

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, *160*(1), 106. https://doi.org/10.1113/jphysiol.1962.sp006837

Hubel, D. H., & Wiesel, T. N. (1963). Shape and arrangement of columns in cat's striate cortex. *The Journal of physiology*, *165*(3), 559. https://doi.org/10.1113/jphysiol.1963.sp007079

Hwang, A. D., Wang, H.-C., & Pomplun, M. (2011). Semantic guidance of eye movements in real-world scenes. *Vision Research*, 51(10), 1192–1205. https://doi.org/10.1016/j.visres.2011.03.010

Indovina, I., & Macaluso, E. (2007). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. *Cerebral Cortex*, *17*(7), 1701-1711. https://doi.org/10.1093/cercor/bhl081

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, 78(7), 2031–2048. https://doi.org/10.3758/s13414-016-1125-4

Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T., & Iriki, A. (2002). Tool-use learning selectively induces expression of brain-derived neurotrophic factor, its receptor trkB, and neurotrophin 3 in the intraparietal multisensory cortex of monkeys. *Cognitive Brain Research*, *14*(1), 3–9. https://doi.org/10.1016/S0926-6410(02)00056-3

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, 73(1), 218–226. https://doi.org/10.1152/jn.1995.73.1.218

Ittelson, W. H. (1951). Size as a Cue to Distance : Static Localization. 64(1), 54–67.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). The principles of psychology.

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*(11), 2463–2475. https://doi.org/10.1093/brain/awg248

James, W., & Drummond, R. (1890). The principles of psychology. *Henry Holt and Company*.

Jolicoeur, P. (1987). A size-congruency effect in memory for visual shape. Memory & Cognition, 15(6), 531–543. https://doi.org/10.3758/BF03198388

Kaas, J. H., Qi, H. X., & Stepniewska, I. (2022). Escaping the nocturnal bottleneck, and the evolution of the dorsal and ventral streams of visual processing in primates. Philosophical Transactions of the Royal Society B, 377(1844), 20210293. https://doi.org/10.1098/rstb.2021.0293

Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, 15(4), 549–559. https://doi.org/10.1016/S0893-6080(02)00048-5

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. Journal of Neuroscience, 17(11), 4302–4311. https://doi.org/10.1523/jneurosci.17-11-04302.1997

Karnath, H. O., Rüter, J., Mandler, A., & Himmelbach, M. (2009). The anatomy of object recognition - Visual form agnosia caused by medial occipitotemporal stroke. *Journal of Neuroscience*, *29*(18), 5854–5862. https://doi.org/10.1523/JNEUROSCI.5192-08.2009

Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kastner, S., Chen, Q., Jeong, S. K., & Mruczek, R. E. B. (2017). A brief comparative review of primate posterior parietal cortex: a novel hypothesis on the human toolmaker. *Neuropsychologia*, *105*, 123–134. https://doi.org/10.1016/j.neuropsychologia.2017.01.034.A

Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: Towards the underlying neural circuitry. *Matters of intelligence: Conceptual structures in cognitive neuroscience*, 115–141. *Springer*. https://doi.org/10.1007/978-94-009-3833-5_5

Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, 33(25), 10235–10242. https://doi.org/10.1523/JNEUROSCI.0983-13.2013

Konkle, T., & Oliva, A. (2011). Canonical Visual Size for Real-World Objects. *Journal of Experimental Psychology: Human Perception and Performance,* 37(1), 23–37. https://doi.org/10.1037/a0020413

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4(1), 47–60. https://doi.org/10.1037/0096-1523.4.1.47

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*(9), 3310-3318. https://doi.org/10.1523/JNEUROSCI.20-09-03310.2000

Kravitz, D.J., Kriegeskorte, N., and Baker, C.I. (2010). High-Level Visual Object Representations Are Constrained by Position. *Cerebral Cortex*, 20:2916-2925. https://doi.org/10.1093/cercor/bhq042

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*, *12*(4), 217–230. https://doi.org/10.1038/nrn3008

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. https://doi.org/10.1016/j.tics.2012.10.011

Kravitz, D.J., Vinson, L.D., and Baker, C.I. (2008). How position dependent is visual object recognition? *Trends in Cognitive Sciences*, 12(3):114-122. https://doi.org/10.1016/j.tics.2007.12.006

Kryklywy, J. H., & Todd, R. M. (2018). Experiential history as a tuning parameter for attention. *Journal of Cognition*, 1(1), 24. https://doi.org/10.5334/joc.25

Kusunoki, M., Gottlieb, J., and Goldberg, M.E. (2000). The lateral intraparietal area as a salience map: the representation of abrupt onset, stimulus motion, and task relevance. *Vision Research*, 40(10), 1459-1468. https://doi.org/10.1016/S0042-6989(99)00212-6

Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science*, 303(5661), 1208–1210. https://doi.org/10.1126/science.1090973

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Leonova, A., Pokorny, J., & Smith, V. C. (2003). Spatial frequency processing in inferred PC- and MC-pathways. *Vision Research*, *43*(20), 2133–2139. https://doi.org/10.1016/S0042-6989(03)00333-X

Leopold, D. A. (2012). Primary visual cortex: awareness and blindsight. Annual review of neuroscience, 35, 91-109. https://doi.org/10.1146/annurev-neuro-062111-150356

Leventhal, A. G., Rodieck, R. W., & Dreher, B. (1981). Retinal ganglion cell classes in the old world monkey: Morphology and central projections. *Science*, *213*(4512), 1142–1144. https://doi.org/10.1126/science.7268423

Li, H.-H., Carrasco, M., & Heeger, D. J. (2015). Deconstructing interocular suppression: Attention and divisive normalization. *PLoS Computational Biology*, 11(10), e1004510. https://doi.org/10.1371/journal.pcbi.1004510

Liu, T. (2019). Feature-based attention: Effects and control. *Current Opinion in Psychology*, 29, 187–192. https://doi.org/10.1016/j.copsyc.2019.03.013

Livingstone, M. S., & Hubel, D. H. (1988). Do the relative mapping densities of the magno- and parvocellular systems vary with eccentricity? *Journal of Neuroscience*, *8*(11), 4334–4339. https://doi.org/10.1523/jneurosci.08-11-04334.1988

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, 145(1), 95–109. https://doi.org/10.1037/xge0000130

Long, B., Yu, C. P., & Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. Proceedings of the National Academy of Sciences of the United States of America, 115(38), E9015–E9024. https://doi.org/10.1073/pnas.1719616115

Lupyan, G. (2017). Changing what you see by changing what you know: The role of attention. *Frontiers in Psychology*, 8, 553. https://doi.org/10.3389/fpsyg.2017.00553

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision*, 11(9), 9–16. https://doi.org/10.1167/11.9.9

Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiati, R. I., Caramazza, A., & Martin, A. (2007). Action-Related Properties Shape Object Representations in the Ventral Stream. *Neuron*, *55*(3), 507–520. https://doi.org/10.1016/j.neuron.2007.07.011

Makovski, T., Hommel, B., & Humphreys, G. (2014). Early and late selection: Effects of load, dilution and salience. *Frontiers in Psychology*, 5, 248. https://doi.org/10.3389/fpsyg.2014.00248

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, 10(2), 4–11. https://doi.org/10.1167/10.2.4

Malcolm, G.L., Rattinger, M., and Shomstein, S. (2016). Intrusive effects of semantic information on visual selective attention. *Attention, Perception, & Psychophysics*, 78:2066-2078. https://doi.org/10.3758/s13414-016-1156-x

Mathews, A., & MacLeod, C. (1986). Discrimination of threat cues without awareness in anxiety states. *Journal of Abnormal Psychology*, 95(2), 131–138. https://doi.org/10.1037/0021-843X.95.2.131

Maunsell, J. H. R., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E., & Noerager, B. D. (1999a). Visual response latencies of magnocellular and

parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience*, *16*(1), 1–14. https://doi.org/10.1017/S0952523899156177

Maunsell, J. H. R., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E., & Noerager, B. D. (1999b). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience*, *16*(1), 1–14. https://doi.org/10.1017/S0952523899156177

Maunsell, J. H. R., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, *10*(10), 3323–3334. https://doi.org/10.1523/jneurosci.10-10-03323.1990

McAnany, J. J., & Alexander, K. R. (2008). Spatial frequencies used in Landolt C orientation judgments: Relation to inferred magnocellular and parvocellular pathways. *Vision Research*, *48*(26), 2615–2624. https://doi.org/10.1016/j.visres.2008.02.012

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4(7), 2–2. https://doi.org/10.1167/4.7.2

Mevorach, C., Tsal, Y., & Humphreys, G. (2014). Low level perceptual, not attentional, processes modulate distractor interference in high perceptual load displays: Evidence from neglect/extinction. *Frontiers in Psychology*, 4, 966. https://doi.org/10.3389/fpsyg.2013.00966

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57–77. https://doi.org/10.1016/0166-4328(82)90081-X

Moore, C. M., & Egeth, H. (1998). How does feature-based attention affect visual processing? *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1296–1310. https://doi.org/10.1037/0096-1523.24.4.1296

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, 6(2), 182–189. https://doi.org/10.1038/nn996

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin & Review*, 23(5), 1316–1340. https://doi.org/10.3758/s13423-015-0982-5

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. Nature Neuroscience, 9(3), 429–434. https://doi.org/10.1038/nn1641

Nah, J.C., Malcolm, G.L., and Shomstein, S. (2021). Task-Irrelevant Semantic Properties of Objects Impinge on Sensory Representations within the Early Visual Cortex. *Cerebral Cortex Communications*, 2:1-11. https://doi.org/10.1093/texcom/tgab049

Nah, J. C., Neppi-Modona, M., Strother, L., Behrmann, M., & Shomstein, S. (2018). Object width modulates object-based attentional selection. *Attention, Perception, & Psychophysics*, 80(6), 1375–1389. https://doi.org/10.3758/s13414-018-1530-y

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*, *10*(5), 360–372. https://doi.org/10.1038/nrn2619

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, 46(5), 614–621. https://doi.org/10.1016/j.visres.2005.08.025

Ni, A. M., Murray, S. O., & Horwitz, G. D. (2014). Object-centered shifts of receptive field positions in Monkey primary visual cortex. Current Biology, 24(14), 1653–1658. https://doi.org/10.1016/j.cub.2014.06.003

Ni, A. M., & Maunsell, J. H. R. (2017). Spatially tuned normalization explains attention modulation variance within neurons. *Journal of Neurophysiology*, 118(3), 1903–1913. https://doi.org/10.1152/jn.00218.2017

Noppeney, U., Price, C. J., Penny, W. D., & Friston, K. J. (2006a). Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex*, *16*(3), 437–445. https://doi.org/10.1093/cercor/bhi123

Noppeney, U., Price, C. J., Penny, W. D., & Friston, K. J. (2006b). Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex*, *16*(3), 437–445. https://doi.org/10.1093/cercor/bhi123

Pandya, D. N., & Yeterian, E. H. (1985). Architecture and connections of cortical association areas. *Association and auditory cortices*, 3-61. https://doi.org/10.1007/978-1-4757-9619-3_1

Perry, V. H., Oehler, R., & Cowey, A. (1984). Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience*, *12*(4), 1101–1123. https://doi.org/10.1016/0306-4522(84)90006-X

Plato. (1973). The Republic (Jowlett, B., translated). *Anchor Books.* (Original work authored ca. 375 B.C.E.)

Pokorny, J., & Smith, V. C. (1997). Psychophysical signatures associated with magnocellular and parvocellular pathway contrast gain: comment. *Journal of the Optical Society of America A*, *14*(9), 2477–2486. https://doi.org/10.1364/josaa.15.002440

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, 2(3), 211–228. https://doi.org/10.1080/02643298508252866

Potter, M. C. (1975). Meaning in visual search. *Science*, 187(4180), 965–966. https://doi.org/10.1126/science.1145183

Ptak, R. (2012). The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. *The Neuroscientist*, 18(5), 502–515. https://doi.org/10.1177/1073858411409051

Retell, J. D., Becker, S. I., & Remington, R. W. (2016). An effective attentional set for a specific colour does not prevent capture by infrequently presented motion distractors. *Quarterly Journal of Experimental Psychology*, 69(7), 1340–1365. https://doi.org/10.1080/17470218.2015.1080738

Revzin, A. M., & Karten, H. (1967). Rostral projections of the optic tectum and the nucleus rotundus in the pigeon. *Brain Research*, *3*(3), 264-276. https://doi.org/10.1016/0006-8993(67)90077-7

Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive psychology*, *24*(4), 502-534. https://doi.org/10.1016/0010-0285(92)90017-V

Rosa, M. G., Tweedale, R., & Elston, G. N. (2000). Visual responses of neurons in the middle temporal area of new world monkeys after lesions of striate cortex. Journal of Neuroscience, 20(14), 5552-5563. https://doi.org/10.1523/JNEUROSCI.20-14-05552.2000

Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of vision*, *12*(4), 14. https://doi.org/10.1167/12.4.14

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual review of vision science*, *2*, 437-457. https://doi.org/10.1146/annurev-vision-082114-035733

Rowe, T. (1996). Coevolution of the mammalian middle ear and neocortex. *Science*, *273*(5275), 651-654. https://doi.org/ 10.1126/science.273.5275.651

Rowe, T. B., & Shepherd, G. M. (2016). Role of ortho-retronasal olfaction in mammalian cortical evolution. *Journal of Comparative Neurology*, *524*(3), 471-495. https://doi.org/10.1002/cne.23802

Sakuraba, S., Sakai, S., Yamanaka, M., Yokosawa, K., and Hirayama, Kazumi. (2012). Does the Human Dorsal Stream Really Process a Category for Tools? *Journal of*

Neuroscience, 32(11), 3949-3953. https://doi.org/10.1523/JNEUROSCI.3973-11.2012

Sali, A. W., Courtney, S. M., & Yantis, S. (2016). Spontaneous fluctuations in the flexible control of covert attention. *Journal of Neuroscience*, 36(2), 445–454.

Sanides, F. (1970). Functional architecture of motor and sensory cortices in primates in the light of a new concept of neocortex evolution. https://doi.org/10.1.1.694.6603

Schulte, P., Alegret, L., Arenillas, I., Arz, J. A., Barton, P. J., Bown, P. R., ... & Willumsen, P. S. (2010). The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. *Science*, *327*(5970), 1214-1218. https://doi.org/10.1126/science.1177265

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. Nature Neuroscience, 14(1), 28–30. https://doi.org/10.1038/nn.2706

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11. https://doi.org/10.1136/jnnp.20.1.11

Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), 586–598. https://doi.org/10.1037/0096-1523.4.4.586

Shellshear, J. L. (1930). I. A study of the arteries of the brain of the spiny anteater (Echidna Aculeata), to illustrate the principles of arterial distribution. *Philosophical Transactions of the Royal Society of London. Series B, Containing*

Papers of a Biological Character, *218*(450-461), 1-36. https://doi.org/10.1098/rstb.1930.0001

Shomstein, S., & Gottlieb, J. (2016). Spatial and non-spatial aspects of visual attention: Interactive cognitive mechanisms and neural underpinnings. *Neuropsychologia*, 92, 9–19. https://doi.org/10.1016/j.neuropsychologia.2016.05.021

Shomstein, S., Malcolm, G. L., & Nah, J. C. (2019). Intrusive effects of task-irrelevant information on visual selective attention: Semantics and size. *Current Opinion in Psychology*, 29, 153–159. https://doi.org/10.1016/j.copsyc.2019.02.008

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of Neuroscience*, 24(47), 10702–10706. https://doi.org/10.1523/JNEUROSCI.2939-04.2004

Silson, E. H., Chan, A. W. Y., Reynolds, R. C., Kravitz, D. J., & Baker, C. I. (2015). A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *Journal of Neuroscience*, *35*(34), 11921-11935. https://doi.org/10.1523/JNEUROSCI.0137-15.2015

Silson, E. H., Groen, I. I., Kravitz, D. J., & Baker, C. I. (2016). Evaluating the correspondence between face-, scene-, and object-selectivity and retinotopic organization within lateral occipitotemporal cortex. *Journal of vision*, *16*(6), 14. https://doi.org/10.1167/16.6.14

Sisk, C. A., Remington, R. W., & Jiang, Y. V. (2018). The risks of downplaying top-down control. *Journal of Cognition*, 1(1), 23. https://doi.org/10.5334/joc.26

Smith, K. J., Valentino, D. A., & Arruda, J. E. (2003). Rhythmic oscillations in the performance of a sustained attention task. *Journal of Clinical and Experimental Neuropsychology*, 25(4), 561–570. https://doi.org/10.1076/jcen.25.4.561.13869

Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. Nature Neuroscience, 15(4), 540–542. https://doi.org/10.1038/nn.3069

Takeshima, Y., & Gyoba, J. (2013). High-intensity sound increases the size of visually perceived objects. Attention, Perception, & Psychophysics, 75(3), 501–507. https://doi.org/10.3758/s13414-012-0403-z

Tanaka, K., Saito, H. A., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, 66(1), 170–189. https://doi.org/10.1152/jn.1991.66.1.170

Theeuwes, J. (2018a). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition*, 1(1), 29. https://doi.org/10.5334/joc.13

Theeuwes, J. (2018b). Visual selection: Usually fast and automatic; seldom slow and volitional; a reply to commentaries. *Journal of Cognition*, 1(1), 21. https://doi.org/10.5334/joc.32

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature human behaviour*, *3*(8), 827-836. https://doi.org/10.1038/s41562-019-0640-4

Todd, R. M., & Manaligod, M. G. M. (2018). Implicit guidance of attention: The priority state space framework. *Cortex*, 102, 121–138. https://doi.org/10.1016/j.cortex.2017.08.001

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision research*, *32*(7), 1349-1357. https://doi.org/10.1016/0042-6989(92)90227-A

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21(4), 753–759. https://doi.org/10.1016/S0896-6273(00)80592-9

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., Wagenmakers, E. (2019). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *PsyArXiv Preprint*, https://psyarxiv.com/spreb

Vaziri-Pashkam, M., & Xu, Y. (2017). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. *Journal of Neuroscience*, *37*(36), 8767–8782. https://doi.org/10.1523/JNEUROSCI.3392-16.2017

Wang, L., Mruczek, R. E., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral cortex*, *25*(10), 3911-3931. https://doi.org/10.1093/cercor/bhu277

Wässle, H., Grünert, U., Röhrenbeck, J., & Boycott, B. B. (1990). Retinal ganglion cell density and cortical magnification factor in the primate. *Vision research*, *30*(11), 1897-1911. https://doi.org/10.1016/0042-6989(90)90166-I

West, G.L., Anderson, A.K., Bedwell, J.S., and Pratt, J. (2010). Red Diffuse Light Suppresses the Accelerated Perception of Fear. *Psychological Science*, 21(7), 992-999. https://doi.org/10.1177/0956797610371966

Wiesel, T. N., & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, *29*(6), 1115–1156. https://doi.org/10.1152/jn.1966.29.6.1115

Wolfe, J.M. (2006). *Sensation & Perception*. Sunderland, Mass: Sinauer Associates.

Wolfe, J.M. (2018). Everything is foreseen, yet free will is given (Mishna Avot 3:15). *Journal of Cognition*, 1(1), 22. https://doi.org/10.5334/joc.27

Wolfe, J.M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28, 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601–621. https://doi.org/10.1037/0096-1523.10.5.601

Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. *Vision Research*, 44(12), 1375-1387. https://doi.org/10.1016/j.visres.2003.12.016

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, 24(1), 3–47. https://doi.org/10.1080/02699930903205698

Zhou, Y. di, & Fuster, J. M. (2000). Visuo-tactile cross-modal associations in cortical somatosensory cells. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(17), 9777–9782. https://doi.org/10.1073/pnas.97.17.9777