
Electronic Thesis and Dissertation Repository

6-25-2019 11:00 AM

The Origins and Development of Visual Categorization

Laura Cabral, *The University of Western Ontario*

Supervisor: Rhodri Cusack, *Trinity College, Dublin*

Joint Supervisor: Jody Culham, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology

© Laura Cabral 2019

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Cognitive Neuroscience Commons](#), [Computational Neuroscience Commons](#), [Developmental Neuroscience Commons](#), and the [Developmental Psychology Commons](#)

Recommended Citation

Cabral, Laura, "The Origins and Development of Visual Categorization" (2019). *Electronic Thesis and Dissertation Repository*. 6292.

<https://ir.lib.uwo.ca/etd/6292>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Forming categories is a core part of human cognition, allowing us to quickly make inferences about our environment. This thesis investigated some of the major theoretical interpretations surrounding the neural basis of visual category development. In adults, there are category-selective regions (e.g. in ventral temporal cortex) and networks (which include regions outside traditional visual regions—e.g. the amygdala) that support visual categorization. While there has been extensive behavioural work investigating visual categorization in infants, the neural sequence of development remains poorly understood. Based on behavioral experiments, one view holds that infants are initially using subcortical structures to recognize faces. Indeed, it has been proposed that the subcortical pathway remains active for rapid face detection in adults. In order to test this in adults, I exploited the nasal-temporal asymmetry of the proposed retinocollicular pathway to see if preferentially presenting stimuli to the nasal hemiretina resulted in a fast face detection advantage when contrasted with presentations to the temporal hemiretina. Across four experiments, I failed to find any evidence of a subcortical advantage but still found that a rapid, coarse pathway exists. Therefore, I moved to investigate the development of the cortical visual categorization regions in the ventral temporal cortex (VTC). I characterised the maturity of the face, place and tool regions found in the VTC, looking at the long-range connectivity in 1-9 month-old infants using MRI tractography and a linear discriminant classifier. The face and place regions showed adult-like connectivity throughout infancy, but the tool-network underwent significant maturation until 9 months. Finally, given this maturity of face and place regions in early infancy, I decided to test whether the organization of the VTC was related to the sequence of categories infants acquire. I used language age of acquisition measurements, determining that infants produce significantly more animate than inanimate words up until 29-months, in line with the animacy distinction in the VTC. My work demonstrates the surprising role and maturity of the cortical regions and networks involved in visual categorization. My thesis develops new methods for studying the infant brain and underscores the utility of publicly available data when studying development.

Lay Abstract

Forming categories is a core part of the human experience. Categorization allows us to recognize people, places and objects. This thesis investigated brain areas involved in visual category development. In adults, there are category-selective brain regions (e.g. in ventral temporal cortex) and networks (which include regions outside traditional visual regions—e.g. the amygdala) that support visual categorization. Behavioural work has shown infants can form categories (e.g. of cats, dogs or faces), but the brain areas infants are using to process visual categories are poorly understood. Based on behavioral experiments, one view holds that infants are initially using subcortical structures to recognize faces. Indeed, it has been proposed that the subcortical pathway infants have been thought to use remains active for rapid face detection in adults. In order to test this in adults, I exploited the nasal-temporal asymmetry of the proposed subcortical pathway to see if preferentially presenting stimuli to the nasal hemiretina resulted in a fast face detection advantage when contrasted with presentations to the temporal hemiretina. Across four experiments, I failed to find any evidence of a subcortical advantage but still found that a rapid, coarse pathway exists. Therefore, I moved to investigate the development of the cortical visual categorization regions in the ventral temporal cortex (VTC). I characterised the maturity of the face, place and tool regions found in the VTC, looking at the long-range white matter, structural connectivity in 1-9 month-old infants. The brain regions selective for faces and places showed adult-like connectivity throughout infancy, but the tool-network underwent significant maturation until 9 months. Finally, given this maturity of face and place regions in early infancy, I decided to test whether the organization of the VTC was related to the sequence of categories infants acquire. I used language age of acquisition measurements, determining that infants produce significantly more animate than inanimate words up until 29-months, in line with the animacy distinction in the VTC. My work demonstrates the surprising role and maturity of the cortical regions and brain networks involved in visual categorization.

Keywords

Face processing, development, infant, nasal-temporal asymmetry, object processing, place processing, ventral temporal cortex, machine learning, large scale databases, category-specific processing.

Co-Authorship Statement

In Chapter 2, Bobby Stojanoski and Rhodri Cusack contributed to the work. Both Bobby Stojanoski and Rhodri Cusack offered helpful experimental design and programming expertise. Rhodri Cusack also contributed to data analysis and gave advice on writing the manuscript. This paper is under review. The citation is:

Cabral, L., Stojanoski, B., Cusack, R. (Submitted). Rapid and coarse face detection in cortex. *Journal of Attention Perception and Psychophysics*.

In Chapter 3, Leire Zubiaurre, Conor Wild, and Annika Linke contributed to data collection. Leire Zubiaurre also contributed diffusion methods expertise. Rhodri Cusack contributed to the data analysis and writing of the manuscript. All co-authors provided comments on the manuscript. The paper has been submitted. The citation is:

Cabral, L., Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (Submitted). Category-Selective visual regions have distinctive connectivity in early infancy. *Nature Communications Biology*.

In Chapter 4, Kristen Turner and Rhodri Cusack contributed to the data analysis. Rhodri also offered helpful comments on the writing of the manuscript. The manuscript is in prep. The citation is:

Cabral, L., Turner, K., Cusack, R. (In Prep). The animacy distinction: Vocabulary development in infancy

Acknowledgments

I would like to thank my supervisors, Rhodri Cusack and Jody Culham for their helpful advice and comments, not only throughout my dissertation writing but throughout my PhD. I have learned a lot from both of them. I would like to thank all of the members of the Cusack and Culham labs for support, meaningful discussions, and helpful programming advice. Special thanks goes to Jacob Matthews and Annika Linke for always being willing to help in any way they could.

Furthermore, I would like to thank my family. Without all of you, none of this would have been possible. Thank you for giving me helpful advice regarding our field (Uncle Steve), letting me borrow your thesis writing schedule (Heather), letting me stay at your house between conferences (Uncle Dave), and touring me around DC (Evelyn). Next, my immediate family has offered me more support than I could have asked for. Thank you for encouraging me, even when you did not understand what I was doing. Finally, I would like to thank my grandmother, who will not see me get this degree but whose support meant everything to me.

Table of Contents

Abstract	ii
Lay Abstract	iii
Co-Authorship Statement.....	v
Acknowledgments.....	vi
Table of Contents	vii
List of Tables	x
List of Figures	xi
List of Appendices	xii
Chapter 1	1
1 General Introduction	1
1.1 Introduction to visual categorization.....	1
1.2 References	11
Chapter 2.....	18
2 Rapid and coarse face detection: with a lack of nasal- temporal asymmetry	18
2.1 Introduction.....	18
2.2 Experiment 1	21
2.2.1 Methods.....	21
2.2.2 Results.....	25
2.2.3 Interim Discussion	28
2.3 Experiment 2	29
2.3.1 Methods.....	29
2.3.2 Results	30
2.3.3 Interim Discussion	30
2.4 Experiment 3	31

2.4.1	Methods.....	31
2.4.2	Results.....	32
2.4.3	Interim Discussion	34
2.5	Experiment 4.....	35
2.5.1	Methods.....	35
2.5.2	Results.....	36
2.5.3	Interim Discussion	37
2.6	General Discussion	37
2.7	Conclusions.....	41
2.8	References.....	42
Chapter 3	48
3	Category-selective visual regions have a distinct signature of connectivity early in infancy.....	48
3.1	Introduction.....	48
3.2	Results.....	49
3.3	Discussion.....	55
3.4	Methods.....	58
3.4.1	Data Acquisition	58
3.4.2	Participants.....	59
3.4.3	Preprocessing	59
3.4.4	Human Connectome Project	60
3.4.5	Tractography and Classification	61
3.5	References.....	63
Chapter 4	67
4	The animacy distinction in the ventral temporal cortex: probing development using vocabulary	67
4.1	Introduction.....	67

4.2 Methods.....	70
4.3 Results.....	72
4.4 Discussion.....	77
4.5 References.....	81
Chapter 5.....	84
5 Discussion.....	84
5.1 General Discussion and Conclusions.....	84
5.2 References.....	92
Appendices.....	95
Curriculum Vitae.....	99

List of Tables

<i>Table 1.</i> Proportion of words known for animate and inanimate categories	73
<i>Table 2.</i> Wilcoxon sum rank tests comparing animate and inanimate categories	75

List of Figures

<i>Figure 1.</i> Exemplar stimuli from the four rapid face experiments.	23
<i>Figure 2.</i> A schematic illustrating the experimental configuration for rapid face experiments.	24
<i>Figure 3.</i> Mean reaction times for the fastest 10% and slowest 50% of trials.	26
<i>Figure 4.</i> Proportion of trials correct, for the fastest 10% and slowest 50% of reaction times.	27
<i>Figure 5.</i> The difference in accuracy between the temporal and nasal hemiretina.	28
<i>Figure 6.</i> Power of the faces and houses at each spatial frequency (cycles/image).	36
<i>Figure 7.</i> Regions used for classification and group overlays of the voxels selected by the classifiers.....	51
<i>Figure 8.</i> Voxel classification performance for infants and adults.....	53
<i>Figure 9.</i> The relationship between the age of participants (14 adults and 11 infants) and classification accuracy (d-prime) with best-fit lines.	54
<i>Figure 10.</i> Proportion of words produced for animate (blue) and inanimate (red) categories.	74
<i>Figure 11.</i> Scatter plot comparing the mean AoA from the parental report measure to the mean AoA obtained from the retrospective self-report measure.	76
<i>Figure 12.</i> Cumulative histogram for the proportion of words produced (parental report) or learned (retrospective self-report).	77

List of Appendices

Appendix A: Regions Driving Connectivity.....	95
Appendix B: Ethics Approval.....	96

Chapter 1

1 General Introduction

1.1 Introduction to visual categorization

Visual Categorization

Being able to identify visual categories is an important part of the human experience. For example, categorization allows us to deduce the properties of a stimulus based on our prior experience with similar things (Robinson, Best, Deng, & Sloutsky, 2012; Sloutsky, 2010). The neural basis of visual categorization has been well studied in adults, where it occurs in the ventral visual stream, a pathway specialized for vision for perception (Goodale & Milner, 1992; Goodale, 2011; Milner & Goodale, 2008). Input to the ventral visual stream begins at the retina, which passes information through the optic nerve to subcortical structures including the lateral geniculate nucleus, which in turn transfers information to primary visual cortex in the occipital lobe (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013). The stimulus is then categorized as input descends through the temporal lobe.

Outside of the temporal lobe, there are many other regions, with reciprocal connections, that aid in our representation of visual stimuli and category development, and definitions of the ventral stream have been expanded to include these regions (Kravitz et al., 2013). For example, regions in the frontal lobes, the ventrolateral prefrontal cortex and the orbitofrontal cortex, may be responsible for the top down processing that helps to assign verbal labels (Kravitz et al., 2013; Robinson et al., 2012). Subcortical regions also make unique contributions-- the amygdala can aid in emotional processing (Phelps & LeDoux, 2005).

Additionally, subcortical structures have been proposed to aid in rapid and low spatial frequency detection of particular visual categories (Vuilleumier, Armony, Driver, & Dolan, 2003). An alternative pathway from the retina, to the superior colliculus, through the pulvinar nucleus and terminating in the amygdala--the retinocollicular pathway--has

been proposed to detect faces (Johnson, 2005). Additional research has demonstrated how this pathway could interact with other structures that process faces—the amygdala has been shown to have functional connectivity with regions in the temporal lobe that aid in face processing (Mende-Siedlecki, Verosky, Turk-Browne, & Todorov, 2013).

Finally, the ventral visual stream has many connections with the dorsal visual stream, which instead of being specialized for vision for perception is specialized for vision for action (Goodale & Milner, 1992). Rather than categorizing visual stimuli found in the environment, the dorsal visual stream represents them in absolute spatial coordinates to allow for accurate grip scaling, obstacle avoidance and other actions (Almeida, Fintzi, & Mahon, 2013; Goodale & Milner, 1992; Goodale, 2011; Mahon et al., 2007; Milner & Goodale, 2008). The dorsal stream interacts with the ventral stream to allow for accurate actions to be completed. For example, the ventral stream contributes helpful category-level information to the dorsal stream. When we see an object in the environment, the ventral stream can pass category-level information to the dorsal stream about the weight, texture, and function of the object, allowing for accurate actions to take place (Almeida et al., 2013; Cant & Goodale, 2007; Goodale & Milner, 1992; Goodale, 2011). For example, we know to grasp a hammer on the handle and not the functional end of the tool.

Organization of the Ventral Temporal Cortex

Studies of categorization have largely focused on the ventral temporal cortex (VTC), as it contains representations of both superordinate and basic level categories. The largest scale of organization of the VTC is demarcated by the mid-fusiform sulcus (Grill-Spector & Weiner, 2014; Weiner et al., 2014). The mid-fusiform sulcus bisects the VTC, separating neurons on the lateral side, which respond to animate stimuli from those on the more medial side, which respond to inanimate stimuli, while each have their own distinct cytoarchitectonic and structural connectivity profile (Grill-Spector & Weiner, 2014; Konkle & Caramazza, 2013; Saygin & Kanwisher, 2014; Saygin et al., 2011). Other stimulus properties, such as real-world object size and eccentricity, are organized around the MFS. Small stimuli, with a more foveal representation, tend to be represented on the lateral side of the MFS, while large visual stimuli tend to be represented, with greater

eccentricity, on the medial side of the MFS (Konkle & Oliva, 2012; Hasson, Levy, Behrmann, Hender, & Malach, 2002). Of course, these general organizing principles are not absolute. On the medial side of the MFS, inanimate categories can be further subdivided into large and small stimuli (Konkle & Caramazza, 2013).

These organizing principles can be seen in multi-voxel pattern analysis methods, which probe distributed representation, with inanimate stimuli, plants, tools, places and other manmade objects forming one cluster, while animate stimuli, faces, body parts, and animals formed another (Kriegeskorte et al., 2008). A continuous semantic distribution across the VTC has also been found, with similar items in semantic space being grouped together (Huth, Nishimoto, Vu, & Gallant, 2012).

Adult Processing of Basic Categories

In addition to the superordinate level organization found in the VTC, as mentioned above, there are regions in the VTC in which basic level categories are strongly represented, contrasting with the partial and retinotopic organization of primary visual cortex in the occipital lobe (Grill-Spector & Weiner, 2014; Kravitz et al., 2013). Having functions restricted to particular regions allows for a clustering of neurons that complete the same task, increasing computational efficiency, as a series of long range connections would waste valuable resources (Grill-Spector & Weiner, 2014; Kravitz et al., 2013). First, a region that was specific for faces was found; this region was termed the fusiform face area (Kanwisher, McDermott, & Chun, 1996; Kanwisher, McDermott, & Chun, 1997) Subsequently, a region that responded specifically to places over other visual stimuli was found (Epstein, Harris, Stanley, & Kanwisher, 1999; Epstein & Kanwisher, 1998). Researchers then began to see if there were regions that reflected other visual categories, such as body parts (Downing, Jiang, Shuman, & Kanwisher, 2001). An area in the VTC was also found to respond to tools and was termed the fusiform tool area (Almeida et al., 2013; Chao, Haxby, & Martin, 1999; Kersey, Clark, Lussier, Mahon, & Cantlon, 2016).

Adult networks

Regions in the VTC do not process stimuli alone, but are part of networks where different brain regions come together to make up the cross-modal, motoric and affective associations we associate with particular categories (Binder, Desai, Graves, & Conant, 2009; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Patterson, Nestor, & Rogers, 2007). As each category has unique associations, category-selective regions in the VTC have been shown to have a unique signature of connectivity with the rest of the brain (Saygin & Kanwisher, 2014; Saygin et al., 2011).

The distributed nature of the face processing network has been well studied (Haxby, Hoffman, & Gobbini, 2000; Moeller, Freiwald, & Tsao, 2008; Tsao & Livingstone, 2008). The most prominent region is the fusiform face area (FFA), which is thought to be responsible for processing the identity of a face (Duchaine & Yovel, 2015; Haxby et al., 2000; Kanwisher & Yovel, 2006; Kanwisher et al., 1997). Outside of the VTC, other regions include the occipital face area, thought to represent the initial processing of facial features, and the superior temporal sulcus, which is thought to be responsible for processing the movement within a face (e.g. lips, moving eyebrows) (Haxby et al., 2000). Frontal regions are thought to also contribute to the top down modulation of face processing, including the dorsolateral prefrontal cortex (Kravitz et al., 2013).

For fast face processing, selective for low spatial frequency stimuli, subcortical regions have also been proposed including the regions that belong to the retinocollicular pathway (Johnson, 2005; Mende-Siedlecki et al., 2013; Vuilleumier et al., 2003). Evidence from this pathway comes from work that has demonstrated that the retinocollicular pathway becomes stronger in blindsight than in age matched controls (Tamietto, Pullens, de Gelder, Weiskrantz, & Goebel, 2012).

The place processing network has also been well studied and has many connections to other areas of the brain that aid in navigation. The VTC hub region for the place processing network is the parahippocampal place area, which responds preferentially to places over other stimuli (Epstein, 2008; Epstein et al., 1999; Kanwisher et al., 1996; Kanwisher & Yovel, 2006). The parahippocampal place area is important when relating

things in space to form a coherent representation of a scene. It is thought to be responsible for the overall representation of objects in the environment but not the number of objects within that scene, and responds strongly when learning about spatial relationships in a navigational context (Kravitz, Saleem, Baker, & Mishkin, 2011; Park, Brady, Greene, & Oliva, 2011). Other areas in the place processing network include the retrosplenial cortex (ROC), which is thought to be responsible for processing navigation and context within an environment (Kravitz et al., 2011). It also responds strongly to landmarks (Kravitz et al., 2011). A nearby region also involved in spatial navigation, but with distinct cytoarchitectonics is the posterior cingulate cortex (PCC). The PCC is thought to be responsible for the location of stimuli in the environment. As it is sensitive to shifts in attention within a scene, it is thought to be responsible for transforming stimulus locations from egocentric to allocentric coordinates (Kravitz et al., 2011). As in other category-specific networks, frontal regions play a role, in this case there is a tight connection with the lateral prefrontal cortex, thought to be responsible for the top down control of eye movements. This network also has extensive links to premotor areas, which aid in orienting and navigation. Finally, the hippocampus plays an important role in navigation, as it contains place cells that fire whenever we place ourselves within a familiar environment (O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1979). The exact role of the hippocampus, and whether it is specialized for pattern separation over pattern completion, continues to be a topic for current research (Bakker, Kirwan, Miller, & Stark, 2008; Rolls, 2013; Yassa & Stark, 2011).

A third example of a category-specific network is the network thought to process objects and tools. The lateral occipital complex (LOC) responds preferentially to objects and tools over other categories, while other regions extend further into the ventral stream with the fusiform tool area, which has been found to be selective for tools (Chao et al., 1999; Cichy, Chen, & Haynes, 2011; Kourtzi & Kanwisher, 2000). Regions that focus on color and texture also make important contributions, such as V4, the inferior occipital gyrus, and the collateral sulcus (Cant & Goodale, 2007). Areas in orbitofrontal cortex may aid in top down tool selection (Bar et al., 2006). As objects are highly variable, some which have utility as tools and some which remain passively perceived and rarely acted on (i.e.

decorations), the exact regions for object responsivity remain a subject for future research.

However, the tool processing network has many reciprocal connections with the dorsal visual stream, which as discussed earlier, is specialized for action (Goodale & Milner, 1992). Extensive connections with the dorsal stream are needed to use tools properly (Almeida et al., 2013; Mahon et al., 2007). As discussed, the ventral stream is needed to identify a tool, but the dorsal visual stream is needed to act on these objects. As such, there is interconnectivity with the superior parietal occipital complex, thought to be responsible for reaching, and the human anterior intraparietal sulcus, thought to be responsible for grasping (Goodale, 2011; Kravitz et al., 2011; Milner & Goodale, 2008). Neuropsychology has been particularly productive in elucidating dissociations and interactions between the ventral and dorsal stream, and it continues to explore this relationship.

Disrupted Connectivity in Adults

As discussed above, networks of regions form the rich motoric and affective associations characteristic of rich semantic categories, leading to distinct connectivity profiles for each category-specific region (Binder, Desai, Graves, & Conant, 2009; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Patterson, Nestor, & Rogers, 2007; Saygin & Kanwisher, 2014). Disrupting connectivity between regions is linked with disordered functioning in many health conditions. For example, abnormalities in structural connectivity are associated with brain damage during concussions (Manning et al., 2019); sometimes the deficits from disrupted connectivity are more predictive of disorders than standard clinical measures in mental health (Whitfield-Gabrieli et al., 2016). In congenital prosopagnosia, a disorder where individuals cannot recognize faces from birth, connectivity is often disrupted, where many groups often have normal functioning of the FFA and other face selective structures (Cook & Biotti, 2016). Indeed, in older adults, reduced connectivity, specifically in the inferior frontal occipital fasciculus, a tract that stretches from frontal regions to occipital regions, has been found to degrade in older adults, with the amount of degradation having a strong correlation with performance on a

face detection task (Thomas et al., 2008). Individuals with autism have also been found to have abnormal connectivity in their face processing network, where people who are diagnosed with autism display abnormal behaviour surrounding faces, as measured with eye tracking (Nomi & Uddin, 2015).

Changes in connectivity are not limited to clinical disorders. Experience can also modulate connectivity. For example, increases in connectivity have also been observed with object use—these have been shown to occur in juggling and piano playing (Bengtsson et al., 2005; Scholz, Klein, Behrens, & Johansen-Berg, 2009). These findings have implications for all of the category-specific networks described above. If the regions within the networks, or the white matter connectivity itself, does not develop normally, it follows that there would be deficits, or reorganization, in category-level processing, whether these deficits are clinical or subclinical.

Development of categorical processing

While there has been substantial progress in understanding the neural basis of adult visual categorization, its developmental origins are far from clear. Some research, in monkeys, has demonstrated that experience is necessary for the formation of face specific cortical regions (Arcaro, Schade, Vincent, Ponce, & Livingstone, 2017). However, behaviourally, infants demonstrate early categorical processing, immediately after birth, (Goren, Sarty, & Wu, 1975); specifically, infants will orient preferentially to faces over control stimuli shortly after they are born. At 3-4 months old, infants are able to form visual, perceptual categories (e.g. cats vs. dogs) (Eimas & Quinn, 1994; Quinn, Eimas, & Rosenkrantz, 1993). If they are presented with a series of cats, they will preferentially orient towards a subsequently presented dog, which shows that they can form a stable representation of the distribution of perceptual features of the cat stimuli (Sloutsky & Fisher, 2004; Wilcox, 1999) and detect deviations from the category.

It is not clear what neural pathways infants use for this early categorical processing. It has been suggested they are using the retinocollicular pathway for the face preference they demonstrate early in development (Johnson, 2005; Johnson, Senju, & Tomalski, 2015). This pathway has been described to be involved in ‘quick and dirty’ visual processing,

facilitating fast face detection, especially with low frequency stimuli (Johnson et al., 2015). Researchers have hypothesized that subcortical structures were responsible for infant face processing capabilities because cortex took longer to myelinate and develop (Deoni et al., 2011; Johnson, 2005). However, to my knowledge, only one study has reported an effect ascribed to the retinocollicular pathway in infants (Simion, Valenza, Umiltà, & Barba, 1998). The study used ‘Johnson faces’, a circle for the face with a black square for each eye and one for the mouth, as the target stimulus, and an upside-down version of the eye and mouth configuration as the distractor. In order to target the retinocollicular pathway, the nasal-temporal asymmetry of the pathway was exploited, where the nasal hemiretina has more connections to the retinocollicular pathway. Following this logic, when the researchers found preferential orienting, only when the stimuli were presented to the nasal hemiretina, they concluded that subcortical structures were responsible for infant face processing capabilities. They went as far as to propose that this pathway remains active in the adult brain (Johnson, 2005). However, this is not without controversy, as there has been debate about which structures are responsible for infant and child face processing capabilities, with others suggesting that cortical structures were responsible. The first study to demonstrate and localize cortical involvement in infant face processing was conducted with positron emission tomography (PET) using participants that were already undergoing clinical imaging (Tzourio-Mazoyer et al., 2002). As such, the researchers had limited time to conduct an experimental protocol and were only able to present faces and a blinking light as the control stimulus. Thus, from their study, it was not known whether the cortical region that was responsive to the faces over the blinking light was specifically responsive to faces or to all complex stimuli (Tzourio-Mazoyer et al., 2002). Therefore, other researchers began to argue that current technology and methods were limiting what could be discovered (McKone, Crookes, Jeffery, & Dilks, 2012).

Given the claim that the retinocollicular pathway remains active in adults, determining its contribution in adults could give us more insight into the development of infant face processing, and would allow us to refine the methods to measure it, and assess their sensitivity. Therefore, in Chapter 2, I sought to develop a protocol to investigate whether adults are using this pathway to process faces. In adults, this pathway has been

proposed as a ‘quick and dirty’ (i.e. fast and approximate) route for the detection of faces. In order to probe whether this was true, I exploited the nasal-temporal asymmetry of the retinocollicular pathway, where the nasal hemiretina has more connections to the retinocollicular pathway. I presented stimuli to both the nasal and temporal hemiretina, to see if the nasal hemiretina, with its increased connectivity, was more accurate at rapid early detection. If subcortical structures were more accurate for the nasal hemiretina, contrasted with the temporal hemiretina, I could infer that the retinocollicular pathway was making a significant contribution to the face detection task at fast reaction times. However, to preview the results, across four experiments I found that while there was indeed a fast face detection advantage, there was no measurable contribution from the retinocollicular pathway. A further control experiment failed to find an effect of the retinocollicular pathway when manipulating spatial frequency.

After demonstrating the lack of evidence for subcortical involvement in adult rapid face detection, in Chapter 3, I moved my attention to the cortical networks in infants. Specifically, I focused on the VTC hubs that contribute to face, place, and tool processing in adults. Recent evidence had shown that there was responsiveness to faces, places, and objects in the infant VTC as early as 4-6 months (Deen et al., 2017). However, that work did not characterize the maturity of the distinct signature of connectivity of the networks associated with the category selective regions, which is needed for the encoding of the cross-modal, motoric and affective associations characteristic of rich semantic categories (Binder et al., 2009; Huth et al., 2016; Patterson et al., 2007). Using computational techniques from machine learning, it was found that face and place regions had adult-like connectivity throughout infancy, but the tool-network underwent significant maturation until 9 months.

In Chapter 4, after finding maturity in the infant face, place, and eventually tool network, I turned my attention to what factors may shape the organization of the VTC in infancy, or whether VTC organization shapes knowledge acquisition. Specifically, I asked whether the order of categories that infants acquire matches the organization of the VTC. As discussed above, the VTC has a distinct organization surrounding the mid-fusiform sulcus. In order to examine this, I used infant ‘age of acquisition’ (AoA) measurements,

which allowed me to determine when infants were able to produce words associated with each category. I used the Wordbank database, which contains measurements from the Mac-Aurthur Bates Developmental Inventory, a questionnaire parents use to report their child's vocabulary development (Frank, Braginsky, Yurovsky, & Marchman, 2017). I chose to probe the animacy distinction in the VTC, as animacy is one of the first semantic distinctions that infants acquire and one of the last to deteriorate in semantic dementia; finally, it is one of the fundamental organizing principles of the VTC (Grill-Spector & Weiner, 2014; Hodges, Graham, & Patterson, 1995). Animacy may also be very salient to infants, as infants will pay attention to things that are moving in their environment (McKenzie & Day, 1976; Volkman & Dobson, 1976). Based on the general organizing principle in the VTC, I hypothesized that there would be a differential rate of learning between animate and inanimate categories. I hypothesized that infants would acquire animate words earlier than inanimate words, reflecting the importance of animate stimuli in the environment. In line with the hypothesis, infants produced reliably more animate words than inanimate words until 29 months. Taken together this PhD emphasizes the role that cortex plays in category-level processing, while developing new methods, used for the first time in infants, to assess the maturity of connectivity in the infant brain.

1.2 References

- Almeida, J., Fintzi, A. R., & Mahon, B. Z. (2013). Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(9), 2334–2344.
- Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, *20*(10), 1404–1412.
- Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*, *319*(5870), 1640–1642.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Schmidt, A. M., ... Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454.
- Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, *8*(9), 1148–1150.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.
- Cant, J. S., & Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex*, *17*(3), 713–731.
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919.
- Cichy, R. M., Chen, Y., & Haynes, J.-D. (2011). Encoding the identity and location of objects in human LOC. *NeuroImage*, *54*(3), 2297–2307.
- Cook, R., & Biotti, F. (2016). Developmental prosopagnosia. *Current Biology: CB*, *26*(8), R312–R313.
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ... Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*, 13995.

- Deoni, S. C. L., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., ... Murphy, D. G. M. (2011). Mapping infant brain myelination with magnetic resonance imaging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(2), 784–791.
- 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.
- Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*, 393–416.
- Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65*(3), 903–917.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*(10), 388–396.
- Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, *23*(1), 115–125.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601.
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: an open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.
- Goodale, M. A. (2011). Transforming vision into action. *Vision Research*, *51*(13), 1567–1587.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, *15*(8), 536–548.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490.

- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: implications for the organisation of semantic memory. *Memory*, 3(3-4), 463–495.
- Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458.
- Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6), 1210–1224.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews. Neuroscience*, 6(10), 766–774.
- Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neuroscience and Biobehavioral Reviews*, 50, 169–179.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1996). A module for the visual representation of faces. *NeuroImage*. [https://doi.org/10.1016/s1053-8119\(96\)80363-6](https://doi.org/10.1016/s1053-8119(96)80363-6)
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 17(11), 4302–4311.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1476), 2109–2128.
- Kersey, A. J., Clark, T. S., Lussier, C. A., Mahon, B. Z., & Cantlon, J. F. (2016). Development of Tool Representations in the Dorsal and Ventral Visual Object Processing Pathways. *Cerebral Cortex*, 26(7), 3135–3145.
- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *The Journal of Neuroscience: The Official Journal of the*

- Society for Neuroscience*, 33(25), 10235–10242.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, 74(6), 1114–1124.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(9), 3310–3318.
- 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.
- 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.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., ... Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6), 1126–1141.
- 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.
- Manning, K. Y., Llera, A., Dekaban, G. A., Bartha, R., Barreira, C., Brown, A., ... Menon, R. S. (2019). Linked MRI signatures of the brain's acute and persistent response to concussion in female varsity rugby players. *NeuroImage. Clinical*, 21, 101627.
- McKenzie, B. E., & Day, R. H. (1976). Infants' attention to stationary and moving objects at different distances. *Australian Journal of Psychology*, 28(1), 45–51.
- McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: experience is less important than previously believed. *Cognitive Neuropsychology*, 29(1-2), 174–212.
- Mende-Siedlecki, P., Verosky, S. C., Turk-Browne, N. B., & Todorov, A. (2013). Robust selectivity for faces in the human amygdala in the absence of expressions. *Journal of Cognitive Neuroscience*, 25(12), 2086–2106.
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed.

- Neuropsychologia*, 46(3), 774–785.
- Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science*, 320(5881), 1355–1359.
- Nomi, J. S., & Uddin, L. Q. (2015). Face processing in autism spectrum disorders: From brain regions to brain networks. *Neuropsychologia*, 71, 201–216.
- O’Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1), 171–175.
- O’Keefe, J., & Nadel, L. (1979). The hippocampus as a cognitive map. *The Behavioral and Brain Sciences*, 2(4), 487–494.
- Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling Scene Content from Spatial Boundary: Complementary Roles for the Parahippocampal Place Area and Lateral Occipital Complex in Representing Real-World Scenes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(4), 1333–1340.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987.
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48(2), 175–187.
- Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, 22(4), 463–475.
- Robinson, C. W., Best, C. A., Deng, W. S., & Sloutsky, V. M. (2012). The role of words in cognitive tasks: what, when, and how? *Frontiers in Psychology*, 3, 95.
- Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience*, 7, 74.
- Saygin, Z., & Kanwisher, N. (2014). Structural and functional connectivity fingerprints for face, body, scene, and object perception. *Journal of Vision*, 14(10), 603–603.
- Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D. E., & Saxe, R. R. (2011). Anatomical connectivity patterns predict face selectivity in the fusiform

- gyrus. *Nature Neuroscience*, *15*(2), 321–327.
- Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience*, *12*(11), 1370–1371.
- Simion, F., Valenza, E., Umiltà, C., & Barba, B. D. (1998). Preferential orienting to faces in newborns: A temporal--nasal asymmetry. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(5), 1399.
- Sloutsky, V. M. (2010). From Perceptual Categories to Concepts: What Develops? *Cognitive Science*, *34*(7), 1244–1286.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology. General*, *133*(2), 166–188.
- Tamietto, M., Pullens, P., de Gelder, B., Weiskrantz, L., & Goebel, R. (2012). Subcortical connections to human amygdala and changes following destruction of the visual cortex. *Current Biology: CB*, *22*(15), 1449–1455.
- Thomas, C., Moya, L., Avidan, G., Humphreys, K., Jung, K. J., Peterson, M. A., & Behrmann, M. (2008). Reduction in white matter connectivity, revealed by diffusion tensor imaging, may account for age-related changes in face perception. *Journal of Cognitive Neuroscience*, *20*(2), 268–284.
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience*, *31*, 411–437.
- Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *NeuroImage*, *15*(2), 454–461.
- Volkman, F. C., & Dobson, M. V. (1976). Infant responses of ocular fixation to moving visual stimuli. *Journal of Experimental Child Psychology*, *22*(1), 86–99.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*(6), 624–631.
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., ... Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex.

NeuroImage, 84, 453–465.

Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, 21(5), 680–685.

Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, 72(2), 125–166.

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34(10), 515–525.

Chapter 2

2 Rapid and coarse face detection: with a lack of nasal-temporal asymmetry

Humans have structures dedicated to the processing of faces, which include cortical components (e.g. areas in occipital and temporal lobes) and subcortical components (e.g. superior colliculus and amygdala). Although faces are processed more quickly than stimuli from other categories, there is a lack of consensus regarding whether subcortical structures are responsible for rapid face processing. In order to probe this, we exploited the asymmetry in the strength of projections to subcortical structures between the nasal and temporal hemiretina. Participants detected faces from unrecognizable control stimuli and performed the same task for houses. In Experiments 1 and 3, at the fastest reaction times, participants detected faces more accurately than houses. However, there was no benefit of presenting to the subcortical pathway. In Experiment 2, we probed the coarseness of the rapid pathway, making the foil stimuli more similar to faces and houses. This eliminated the rapid detection advantage, suggesting that rapid face processing is limited to coarse representations. In Experiment 4, we sought to determine whether the natural difference between spatial frequencies of faces and houses were driving the effects seen in Experiments 1 and 3. We spatially filtered the faces and houses so that they were matched. Better rapid detection was again found for faces relative to houses, but we found no benefit of preferentially presenting to the subcortical pathway. Taken together, the results of our experiments suggest a coarse rapid detection mechanism, which was not dependent on spatial frequency, with no advantage for presenting preferentially to subcortical structures.

2.1 Introduction

Animals as diverse as fish, birds and sheep can recognize the faces of their conspecifics (Leopold & Rhodes, 2010). In humans there has evolved a network of structures responsible for face processing that facilitates face detection, orientating, and identification (Haxby, Hoffman, & Gobbini, 2000; Mende-Siedlecki & Verosky, 2013; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000). This comprises subcortical

components, including the superior colliculus and amygdala (Mende-Siedlecki & Verosky, 2013; Vuilleumier, Armony, Driver, & Dolan, 2003), and cortical components in the occipital and temporal lobes (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006; Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). These specialized processing mechanisms allow faces to be detected more quickly than objects (Crouzet, Kirchner, & Thorpe, 2010) and result in faces being the first category detected in visual search tasks (Fletcher-Watson, Findlay, Leekam, & Benson, 2008). Detecting faces quickly is thought to be evolutionarily advantageous for both survival and social interaction, from the savannahs of Africa to the office party.

The subcortical route via the retinocollicular pathway to the amygdala is often thought to facilitate ‘quick and dirty’ face detection (Johnson, 2005). It comprises projections from the retina to the superior colliculus, which in turn project to the pulvinar nucleus on the way to the amygdala (Benevento & Standage, 1983; Jones & Burton, 1976; Rafal et al., 2015; Tamietto, Pullens, De Gelder, Weiskrantz, & Goebel, 2012). Evidence that the retinocollicular pathway can process faces comes from blindsight patients, who after extensive damage to visual cortex are still able to detect the emotional content of faces, although they cannot recognize their identity (Tietto & de Gelder, 2010). Similar behavior is found in healthy controls following transcranial magnetic stimulation to the visual cortex; when TMS prevents participants from seeing stimuli, they are still able to recognize the emotional content of the face (Jolij & Lamme, 2005). Furthermore, structures in the retinocollicular pathway are activated by the viewing of neutral and emotional faces, as shown with functional magnetic resonance imaging (fMRI) (Mende-Siedlecki & Verosky, 2013). Functional magnetic resonance imaging has also found that this pathway has a preference for crude, low-spatial frequency information, with greater activation to faces filtered to emphasize low spatial frequencies than high spatial frequencies (Vuilleumier et al., 2003).

Intracranial recordings in epilepsy patients have found that the retinocollicular pathway is fast, with neural firing in the amygdala as quickly as 100-250 ms after the presentation of an emotional face (Sato et al., 2013). Recent intracranial recording from Méndez-Bértolo et al. (2016) has found even faster processing for fearful faces, with firing in the

amygdala recorded 74 ms after stimulus onset. Magnetoencephalography (MEG) data suggests even faster processing with responses to emotional faces detected in just 40 ms (Luo et al., 2010). Supporting this hypothesis, Garvert, Friston, Dolan, & Garrido (2014) used dynamic causal modeling of MEG data to conclude that a model with a subcortical component, containing the pulvinar nucleus and the amygdala, more accurately modeled rapid face processing than a model with a singular cortical process.

It has been proposed that these putative fast face detection mechanisms are not limited to subcortical structures, as there is also evidence of rapid mechanisms within cortical areas, such as the inferior occipital gyrus (Pitcher, Walsh, Yovel, & Aviv, 2007; Sadeh, Podlipsky, Zhdanov, & Yovel, 2010). Specifically, an initial feed-forward wave of firing through cortex could allow for rapid, coarse processing (Cauchoix & Crouzet, 2013; Serre, Oliva, & Poggio, 2007; Vanrullen & Koch, 2001). Electroencephalography (EEG) data from visual cortex can identify responses just 56 ms after stimulus onset (Foxe & Simpson, 2002), and intracranial recordings in epilepsy patients found that the category of image participants were viewing could be decoded from the first 100 ms of response in visual cortex (Liu, Agam, Madsen, & Kreiman, 2009). MEG data suggests occipitotemporal responses to faces in just 100 ms (Liu, Harris, & Kanwisher, 2002). Barragan-Jason, Cauchoix, & Barbeau (2015) have proposed that even the identification of familiar faces has an initial rapid phase, occurring at 140 ms, that depends on coarse visual information, and behavioural responses to familiar faces can be detected in just 180 ms (Visconti di Oleggio Castello & Gobbini, 2015). To formalize how cortex could rapidly detect complex visual objects such as faces in real-world scenes, Thorpe and colleagues (Delorme & Thorpe, 2001; VanRullen, Guyonneau, & Thorpe, 2005) proposed a spike-based model of rapid processing. These models have been supported by recordings from V1 in the macaque and cat (Celebrini, Thorpe, Trotter, & Imbert, 1993; Konig, Engel, Roelfsema, & Singer, 1995; VanRullen et al., 2005).

In summary, although not without its critics, many authors have argued for both subcortical and cortical mechanisms for rapid visual processing of faces. Which one, therefore, dominates rapid face detection in healthy participants? One way to address whether rapid face perception is driven by subcortical structures is to target the

retinocollicular pathway to the amygdala. Presenting stimuli exclusively to the nasal hemiretina preferentially targets the retinocollicular pathway, as the nasal hemiretina contains more fibers projecting to the superior colliculus. Initial evidence for this asymmetry came from tree shrews, cats and macaques (Conley, Lachica, & Casagrande, 1985; Harrison, 2015; Perry & Cowey, 1985; Pollack & Hickey, 1979; Sterling, 1973). fMRI evidence in humans has demonstrated that the superior colliculus displays a temporal nasal asymmetry that is not found for the LGN or V1 (Sylvester, Josephs, Driver, & Rees, 2007). Additionally, behavioral studies have demonstrated that a nasal-temporal asymmetry is reflective of input to the superior colliculus. For example, making stimuli only visible to the S cones, which do not provide input to the superior colliculus, eliminates the benefit of presenting to the nasal hemiretina (Bertini, Leo, & Làdavas, 2008).

Our goals in this study were to establish a paradigm for behaviorally quantifying rapid face detection, and to determine whether presenting preferentially to the retinocollicular pathway resulted in improved rapid face detection. Participants were asked to detect faces from amongst unrecognizable control stimuli that were matched to have the same low-level visual features, as quantified with a model of the early visual system (Stojanoski & Cusack, 2014). To determine whether any rapid detection mechanism was specific to faces, we also tested a control condition, requiring detection of another class of visual object, houses.

2.2 Experiment 1

2.2.1 Methods

To probe rapid face processing, in two blocks, participants performed a face detection task in which they pressed a button as quickly as possible for intact faces, but not for scrambled foil stimuli. In two additional blocks, they were asked to detect houses in a similar manner. In each block, stimuli were presented monocularly, by asking participants to wear an eye patch. This allowed us to target stimuli exclusively to either the nasal or temporal hemiretina. In the right eye, presenting stimuli to the right of fixation targets the nasal hemiretina, while presenting to the left of fixation targets the

temporal hemiretina. The opposite is true in the left eye. Within each block, stimuli were randomized across the nasal and the temporal hemiretinas.

2.2.1.1 Participants

Twenty-four individuals (12 males, 12 females, age range 18-21) were given course credit for participation in Experiment 1. The non-medical ethics board at the University of Western Ontario reviewed and approved the experimental protocol. All participants gave informed consent, reported normal or corrected-to-normal vision, and that they were right handed.

2.2.1.2 Stimuli

Twenty-four face photographs from an online database (http://wiki.cnhc.cmu.edu/Face_Place) and 24 house stimuli, created by Martin, McLean, O'Neil, & Köhler (2013), were used in the study. As the house stimuli had a blurred edge, a custom Matlab script added a blurred edge to the face stimuli, to appear similar by eye. As the house stimuli were greyscale, face stimuli were also altered to be greyscale.

All stimuli were centered in a rectangular area of 4.9 degrees by 4.9 degrees of visual angle. The fixation cross was .5 degrees by .5 degrees. A white background was used throughout the experiment. In all experiments, participants viewed the stimuli in a room with the lights on. To generate the control stimuli, faces and houses were diffeomorphically warped using the procedure described by Stojanoski & Cusack (2014). Foils were unrecognizable as determined by the behavioural ratings in Stojanoski & Cusack (2014) (image 38 on the diffeomorphic continuum). A depiction of the stimuli used in Experiment 1 can be found in Figure 1 A.

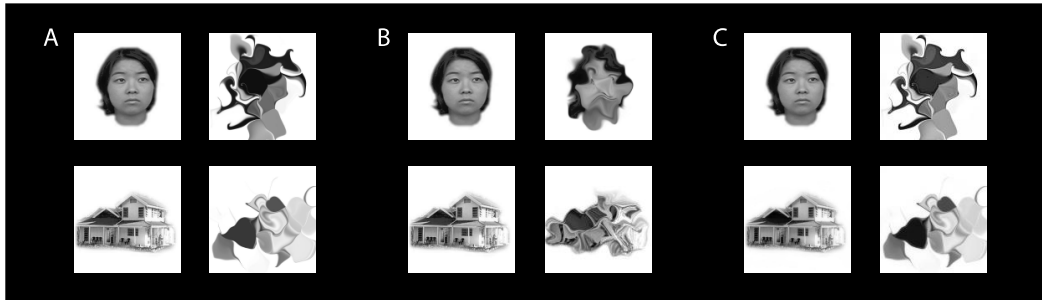


Figure 1. Exemplar stimuli from the four rapid face experiments.

A) Stimuli used in Experiments 1 and 3. Foil stimuli are unrecognizable versions of faces and houses B) In Experiment 2, the foil stimuli were more similar to the faces and houses. C) In Experiment 4, the stimuli that were adjusted for differences in spatial frequency between categories.

2.2.1.3 Procedure

Stimuli were presented on a laptop screen using MATLAB and Psychtoolbox.

Participants wore an eye patch to ensure monocular presentation, placed their heads on a chin rest, and were instructed to maintain fixation. The centre of the screen was directly ahead of the nose. In each experimental block, a black fixation cross was offset by 3.2 cm to the left or right from center in order to put it directly in front of the unpatched eye. This distance was chosen using the mean interpupillary distance scores from the 1988 Anthropometric Army Survey.

In Experiment 1, participants completed two blocks with their left eye unpatched, one that contained only face targets, the other containing house targets, and two similar blocks with their right eye unpatched. Block order was counterbalanced across participants.

In each block, participants were presented with 96 trials comprising two repetitions of 24 target stimuli and their 24 warped counterparts. One repetition was presented to the nasal hemiretina, while the other was presented to the temporal hemiretina. To present to the nasal and temporal visual hemiretina, the stimuli were offset horizontally so that the outer edge of their rectangular bounds was 8 degrees from the center of fixation. Stimuli were presented for duration of 122 ms, with an inter-trial interval of 2505 ms. Participants

were instructed to perform a simple detection task, pressing a key as quickly as possible when they saw an intact face (in the face blocks) or an intact house (in the house blocks). For a schematic of the experimental configuration, please see Figure 2.

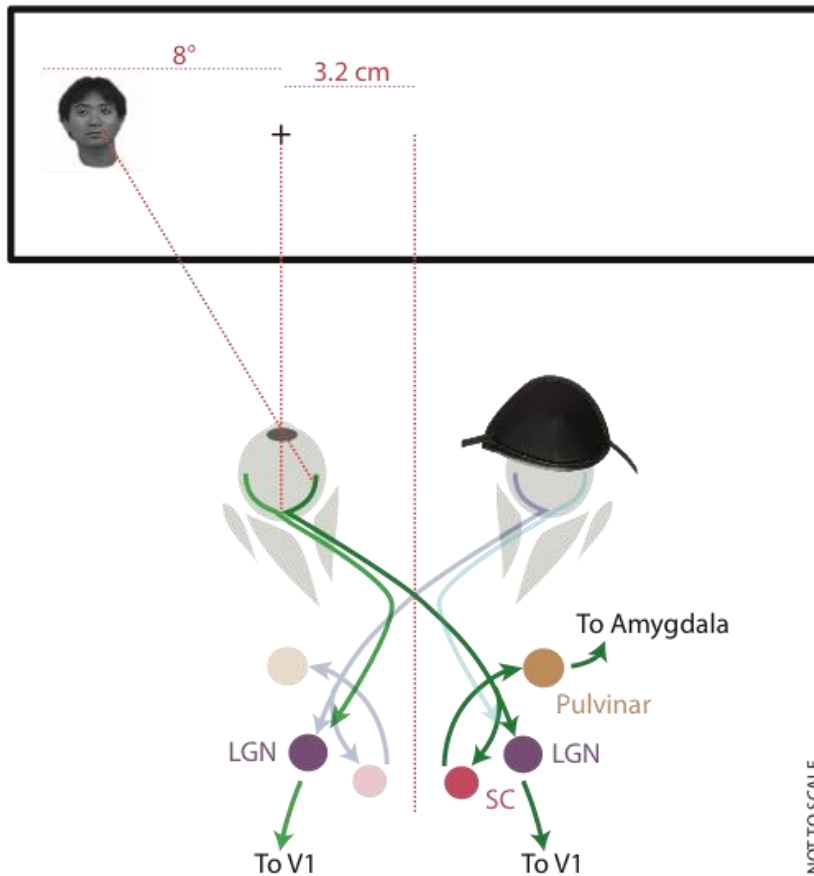


Figure 2. A schematic illustrating the experimental configuration.

Participants wore an eye patch to ensure monocular presentation and to allow the stimuli to be presented exclusively to the nasal or temporal hemiretina. The retinocollicular pathway is depicted with projections from the nasal hemiretina to the superior colliculus, through the pulvinar nucleus, terminating in the amygdala. Weaker projections from the temporal hemiretina to the superior colliculus are not shown.

2.2.1.4 Analysis

In order to quantify rapid processing, we used an analysis strategy similar to Kirchner & Thorpe's (2006) and calculated accuracy for the fastest 10% of responses. All reaction times are relative to stimulus onset. A fast detection mechanism would be expected to improve accuracy on these rapid trials by providing more accurate information to

decision and action areas sooner after stimulus onset. The reaction time (RT) threshold for the fastest 10% of trials was calculated for each participant individually, in order to account for individual differences in overall reaction time. We also expected that faces would be detected more quickly overall. If this is the case, to ensure that the overall difference in reaction time between the faces and houses did not drive the results, we adopted a conservative analysis strategy and determined the face and house reaction thresholds separately. Thus, the fastest 10% of face trials were expected to be even faster than the fastest 10% of house trials.

To determine the contribution of the retinocollicular pathway, we examined whether presenting the stimuli to the nasal or the temporal hemiretina modulated performance. As the nasal hemiretina has more connections to the superior colliculus and thus the retinocollicular pathway, we would expect to see faces more accurately detected than houses, when the stimuli are presented to the nasal hemiretina.

2.2.2 Results

Two participants were excluded for failing to follow the task instructions. Across the remaining participants, mean reaction times for both the fastest 10% and slowest 50% of trials are shown in Fig. 3A. These reaction times include correct responses and false alarms, as both contributed to subsequent accuracy metrics.

To probe rapid mechanisms, analyses were confined to trials with a rapid response, in the fastest 10% of RTs for each category. Participants were able to more accurately detect faces than houses ($F(1,21)=10.41$, $p<0.01$) (Fig. 4 A). This shows that our paradigm is sensitive to rapid, accurate face detection. We then turned to the effect of the retinal hemifield manipulation. There was no overall benefit of presenting stimuli to a particular hemiretina ($F(1,21)= 3.87$, $p=0.062$), suggesting no general role for the retinocollicular pathway in fast visual detection. Furthermore, contrary to what would be expected if the retinocollicular pathway was category selective, and supported rapid face detection, there was no significant stimulus by retinal hemifield interaction ($F(1,21)=0.1$, $p=.755$) (Fig. 5A). In fact, there was a trend for better performance for faces in the temporal hemiretina.

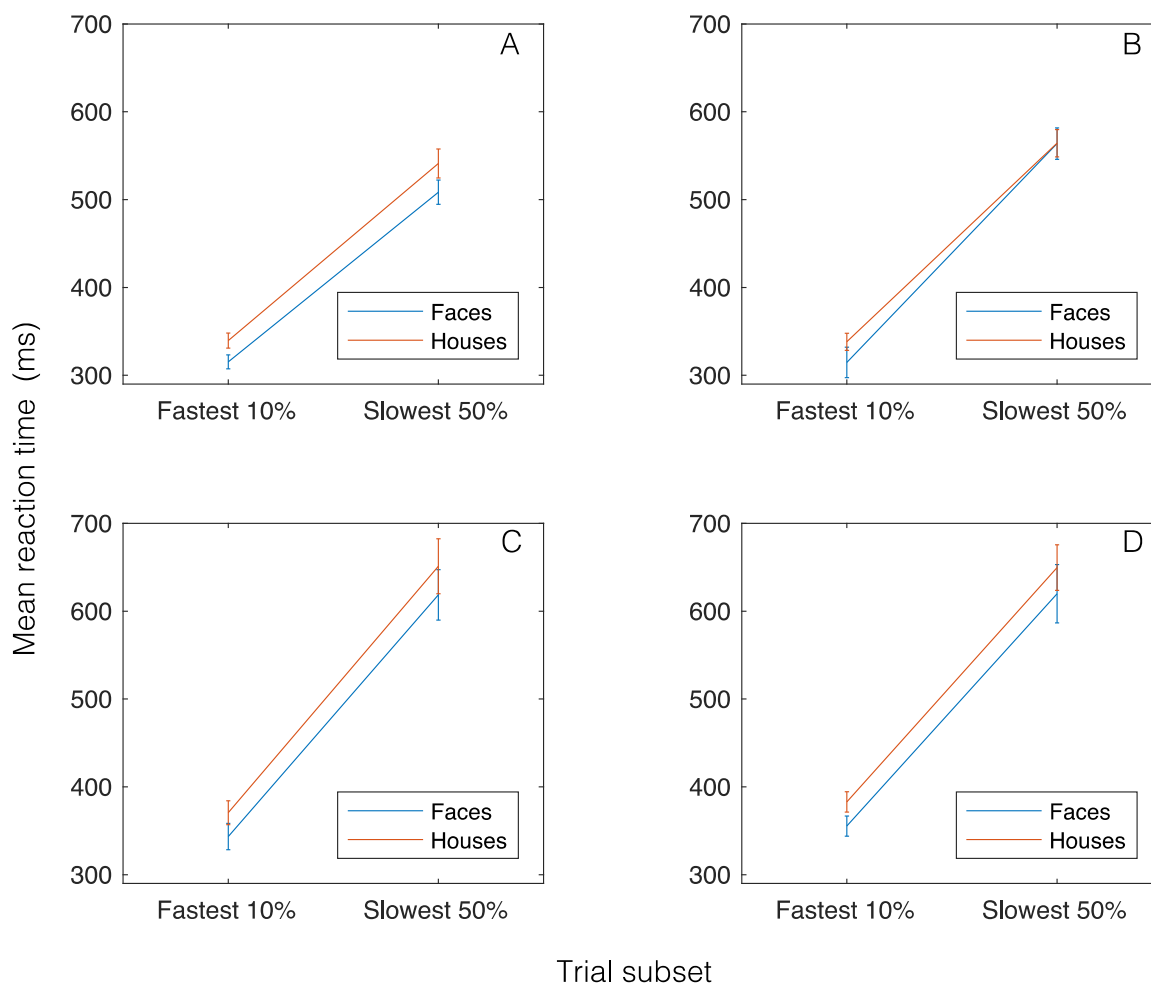


Figure 3. Mean reaction times for the fastest 10% and slowest 50% of trials.

A) In Experiment 1, foil stimuli were unrecognizable versions of faces and houses B) In Experiment 2, faces and houses were more similar to foil stimuli C) In Experiment 3, stimuli were the same as in Experiment 1 D) In Experiment 4, spatial frequency of the faces and houses were matched. Error bars represent plus-or-minus one standard error.

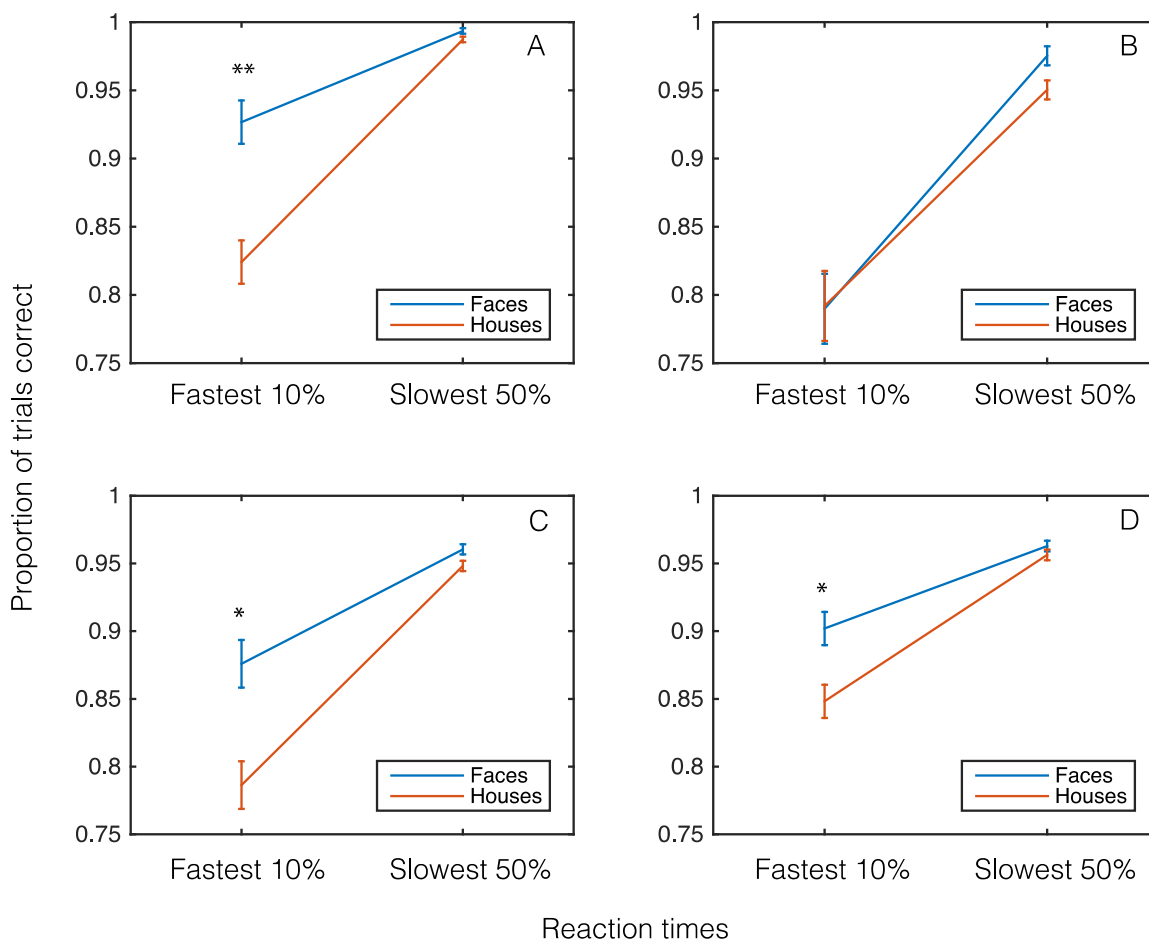


Figure 4. Proportion of trials correct, for the fastest 10% and slowest 50% of reaction times. A) In Experiment 1, faces were detected significantly more accurately than houses at the fastest reaction times. B) In Experiment 2, faces and houses were detected with similar accuracy. C) Experiment 3 replicated the results of Experiment 1. D) In Experiment 4, faces were detected significantly more accurately than houses in the fastest 10% of reaction times. In all experiments, error bars represent +/- the standard error

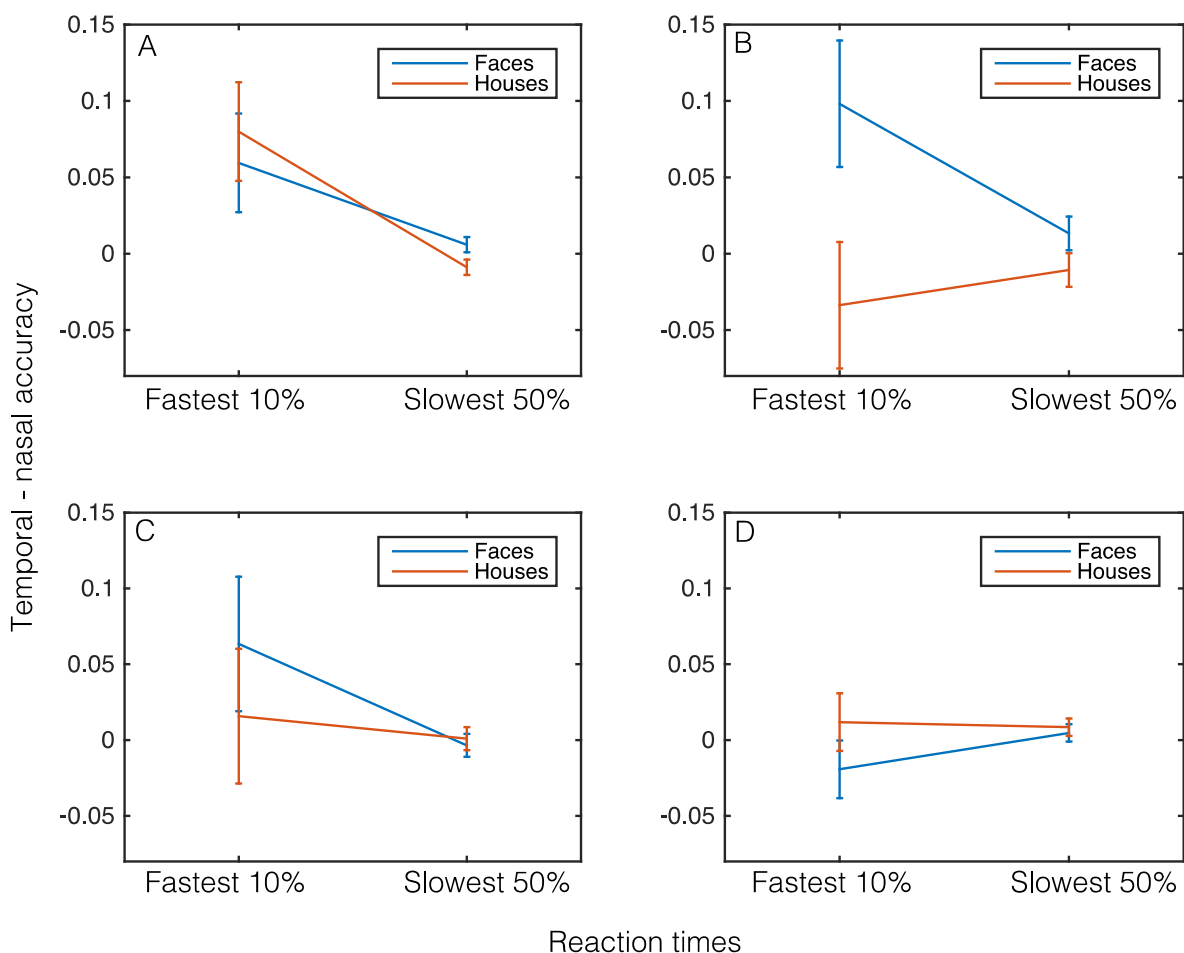


Figure 5. The difference in accuracy between the temporal and nasal hemiretina.

Plotted for the fastest 10% and slowest 50% of RT's for Experiments 1-4 (A-D, respectively).

There were no significant differences between the nasal and temporal hemiretina for the faces and houses in any of the experiments. Error bars represent plus-or-minus one standard error.

2.2.3 Interim Discussion

The results of Experiment 1 demonstrate that there is a rapid route for detecting faces that does not extend to other classes of stimuli (i.e. houses). As there was no benefit for presenting stimuli to the nasal hemiretina, the results of the experiment did not provide any evidence of a role for the retinocollicular pathway in rapid visual detection or rapid face processing. The lack of contribution from the retinocollicular pathway, taken with

the trend for better processing in the temporal hemiretina, suggests that a cortical route could be responsible for the rapid face detection seen in the experiment.

Our next goal was to probe the specificity of the rapid pathway. A key feature of the rapid route discussed in the literature is that it is not just quick, but that it is dirty (i.e. a coarse representation). In an evolutionary context, it might be advantageous for neural structures to obtain extremely quick, coarse representations of the faces in the environment. This route is not thought to be capable of fine discrimination. Thus, the next experiment was designed to probe the precision of the rapid detection mechanism identified in Experiment 1.

2.3 Experiment 2

2.3.1 Methods

In order to examine the precision of the rapid cortical detection route, participants performed the same task as in Experiment 1, but with less warped foil stimuli. These foil stimuli still had some recognizable features of faces and houses. If detection relied on a rapid route, exclusively for faces, it would support the idea that the rapid detection mechanism was capable of precise representations. Otherwise, the rapid detection mechanism might be limited to rapid, coarse judgments.

2.3.1.1 Participants

The same participants who participated in Experiment 1 participated in Experiment 2, and the order in which participants completed the two experiments was counterbalanced. Again, two participants were excluded for failing to follow the task instructions.

2.3.1.2 Stimuli and Procedure

The stimuli and procedure were the same as in Experiment 1, except that the foil images had less warping applied (image 5 in the diffeomorphic continuum). Examples of the stimuli can be found in Fig. 1B.

2.3.2 Results

As in Experiment 1, overall mean reaction times for the fastest 10% and slowest 50% of trials are shown in Fig. 3B. The mean reaction times include both correct responses and false alarms. We used the same analysis procedure as in Experiment 1, with accuracy in the fastest 10% of trials used to assess rapid face detection. When participants were required to make precise judgments, faces were no longer detected reliably more accurately than houses ($F(1,21)=0.08$, $p=.784$) (Fig. 4B). Again, to examine the role of the retinocollicular pathway we compared presentations to the nasal or the temporal hemiretina. At the fastest RTs, there was no significant difference in accuracy between hemiretinas ($F(1,21)=0.29$, $p=.598$). Furthermore, again there was no evidence that faces were detected significantly more accurately than houses in the nasal hemiretina when compared to the temporal hemiretina ($F(1,21)=3.97$, $p=0.059$) (Fig. 5B).

In Experiment 1, we found evidence of a fast face processing mechanism when faces were clearly distinct from foils. In Experiment 2, with a smaller difference between faces and foils, we did not find the same effect. However, it is important to establish whether the effect of the foil manipulation was significant, by directly testing whether the results of the two experiments are significantly different. This comparison showed that performance was significantly more accurate in Experiment 1 than in Experiment 2 ($F(1,21)=6.81$, $p<0.05$). Furthermore, there was a significant interaction between the experiments and stimulus type ($F(1,21)=8.03$, $p<0.05$). This is driven by a greater difference between rapid detection of faces and houses in Experiment 1 than in Experiment 2.

2.3.3 Interim Discussion

When foil stimuli were created with less warping, requiring participants to make fine discriminations, faces were no longer detected more accurately than houses at the fastest reaction times. Again, there was not a significant advantage, or a trend for better performance, when stimuli were preferentially presented to the retinocollicular pathway. The results of this experiment support the idea that rapid detection of faces is limited to

coarse visual characteristics. When taking Experiment 1 and Experiment 2 together, the results support the idea that there is no advantage of presenting to subcortical structures.

One weakness of the current analysis is that the comparisons of the nasal and temporal hemiretina contain half as much data as the collapsed analyses, and perhaps the consequently reduced power that results is responsible for the lack of significance. Thus, we conducted a further experiment, to double the number of subjects for this comparison. Given recent concerns about the reproducibility of results in psychology (Open Science Foundation, 2015) this also affords us the opportunity to test for replication of the other findings from Experiment 1.

2.4 Experiment 3

2.4.1 Methods

Experiment 3 was conducted to ensure that the results from Experiment 1 were generalizable, replicating it in a different group of participants. We sought to combine the participants from Experiment 1 and Experiment 3 into a larger analysis, where we would have increased power to detect differences in performance between the nasal and temporal hemiretina.

2.4.1.1 Participants

Twenty-five self-reported right handed individuals (12 males, 13 females, age range 18-42) participated in Experiments 3 and 4. Twenty-four participants reported normal or corrected to normal vision. One participant did not have corrected to normal vision, their prescription was +0.75 for the right eye and +0.5 for the left eye. Two participants were excluded from the experiment, one because a fire alarm occurred during their experimental session and the other because of technical difficulties that prevented button presses from being recorded.

The participants received \$10 for their participation in the experiment. All participants gave written informed consent. The non-medical ethics board at the University of Western Ontario reviewed and approved the experimental protocol.

2.4.1.2 Stimuli and Procedure

Stimuli were identical to those that were used in Experiment 1. One important change was made to the procedure. In order to gain information about the participants' reaction times in both warped and intact trials, participants were instructed to press two buttons, one for the warped images and another for the intact images. Exemplar images of the stimuli can be found in Fig. 1A.

2.4.2 Results

Reaction times for the fastest 10% and slowest 50% of trials are shown in Fig. 3C. As in Experiments 1 and 2, accuracy in the fastest 10% of trials was examined. In this experiment, we included the data from both the target and foil trials in our analysis. Replicating the findings from Experiment 1, faces were detected significantly more accurately than houses at faster RTs ($F(1,22)=6.24, p<0.05$) (Fig. 4C). Again, when collapsed across faces and houses, no difference in accuracy at fast RTs was found across the nasal and temporal hemiretina ($F(1,22)=1.88, p=.184$). Furthermore, the interaction between the visual field and stimulus class showed that faces were not significantly more accurate than houses in the nasal hemiretina than the temporal hemiretina ($F(1,22)=.19, p=.667$) (Fig. 5C).

In order to test if a difference in response bias was responsible for the difference in accuracy at the fastest reaction times, we also calculated the false alarm and hit rate for the faces and houses. We were able to do this in Experiment 3 because it was a two-button response task, which allowed us to bin all responses by RT. The mean false alarm rate was lower for faces ($M=.21, SE=0.016$) than for houses ($M=.25, SE=0.016$) at the fastest 10% of RT's. The mean hit rate was higher for faces ($M=.91, SE=0.017$) than for houses ($M=.82, SE=0.017$) at the fastest 10% of RT's. A higher hit rate and a lower false alarm rate shows the results were not driven by a response bias and participants were actually better at identifying faces than houses. The higher hit rate and low false alarm rate for faces suggests that participants were not merely responding less carefully to the rapid face trials and that the results were not a result of a speed accuracy trade off.

Although a response bias does not appear to be causing the results in the experiment, it is possible that the effect of hemifield is not being seen because of insufficient power. Therefore, we conducted a further analyses in which we included participants from both Experiment 1 and Experiment 3 yielding $N=45$. When comparing the results from Experiment 1 to Experiment 3, we tested whether the results from the two experiments were significantly different; they were not $F(1,43)=1.23, p=.274$. In both the nasal and the temporal hemiretina, a significant difference in accuracy at fast reaction times was found for face compared with house detection ($t(1,44)=2.03, p<0.05, t(1,44)=3.91, p<0.001$, respectively). This supports the idea that increases in face detection accuracy are not driven exclusively by an increase in performance in the nasal hemiretina, as would be expected if the retinocollicular pathway were responsible.

Further combined analyses from Experiment 1 and Experiment 3 replicated the key results. At the fastest reaction times, faces were detected more accurately than houses ($F(1,44)=16.44, p<0.001$), consistent with the results of previous experiments. In addition, at the fastest RTs, overall performance in the nasal hemiretina was significantly worse than performance in the temporal hemiretina ($F(1,44)=5.74, p<0.05$). With the larger sample, there was still no significant interaction between stimulus and field ($F(1,44)=0.03, p=.862$) as would be expected if a nasal benefit was driving improved face detection.

One criticism of the approach we have taken is that frequentist statistics only allow for the inability or ability to reject the null hypothesis, whereas Bayesian statistics allow us to estimate the probability of null and other models. In order to address this, in our pooled analysis (45 participants over experiments 1 and 3), we conducted a Bayesian Repeated Measures ANOVA with default prior settings in JASP. There was moderate evidence against a field and stimulus interaction ($BF_{10}=4.6$). A difference would be expected between the nasal faces and houses if the retinocollicular pathway was driving the effects.

2.4.3 Interim Discussion

Experiment 3 replicated the results of Experiment 1, generalizing the findings to a different group of participants and a slightly different response procedure. In addition, calculating the false alarm and hit rates allowed us to determine that a response bias was not the cause of our results. A higher hit rate and a lower false alarm rate for faces suggests that increased accuracy is not a result of a speed accuracy trade off. This is further emphasised because faces have a faster mean reaction time than houses.

Combining the results from Experiments 1 and 3 into a single analysis revealed that presentation to the nasal hemiretina led to significantly worse rapid detection of faces and houses. This result is contrary to what would be expected if the nasal hemiretina, and thus the retinocollicular pathway, were driving the results. In addition, when looking at frequentist statistics, both the nasal and the temporal hemiretina show evidence of significantly more accurate face detection at fast reaction times, demonstrating that there is not one hemiretina driving the fast face detection advantage.

Taken together, these results provide support for the idea that a cortical, rather than a subcortical, process is responsible for rapid face detection. However, we want to be clear that the conclusion is based on our inability to reject the null hypothesis over multiple experiments.

Why might presenting to the nasal hemiretina result in reduced detection of visual stimuli? It is possible that reduced performance could be caused by distracting information (i.e. emotional content) being communicated from subcortical structures to cortical structures. At fast RTs, the brain might be only capable of attending only to a subset of information, and emotional content might take precedence over visual categorization, decreasing the accuracy of the nasal hemiretina in Experiments 1 and 3.

In a final experiment we control for a potential low-level visual explanation for the category specificity of the rapid detection mechanism. In our stimulus sets (and more generally, Awasthi, Sowman, Friedman, & Williams, 2013), faces contained lower spatial frequencies than houses. Natural images generally have greater power at lower

frequencies (Burton & Moorhead, 1987) so perhaps we have more rapid mechanisms for low spatial frequencies, that process faces more rapidly. Thus, perhaps spatial frequency, rather than category *per se* is responsible for the category specific rapid detection we observed in Experiments 1 and 3.

2.5 Experiment 4

2.5.1 Methods

In Experiment 4 we repeated Experiment 3, but the face and house stimulus sets filtered so that they had balanced power spectra.

2.5.1.1 Participants

Experiment 4 tested the same participants as Experiment 3, and the order in which they participated was counterbalanced. Again two participants were excluded - one because of a fire drill and the other because of technical difficulties that prevented button presses from being recorded.

2.5.1.2 Stimuli and Procedure

The same stimuli that were used in Experiment 1 and 3 were used in Experiment 4, but with the spatial frequency of the images balanced. Each image was transformed into 2D frequency space using a Fourier transform. Each pixel was then multiplied by a scalar filtering function that depended only on distance from the origin of frequency space. Finally, an inverse Fourier transform was used to return to image space. Houses were filtered to remove high spatial frequency information, and faces were filtered to remove low spatial frequency information. A further processing stage was applied, to remove a visually salient artefact, which was the bleeding of images into the background surrounding them. All voxels outside of each object in the original image (i.e., that were exactly background color) were reset to the background color after filtering. This led to a slight residual mismatch in the resulting frequency spectra, which can be seen in the original and final frequency spectra, shown in Fig. 6. Exemplar images can be found in Fig. 1C. All other aspects of the experiment were the same as in Experiment 3.

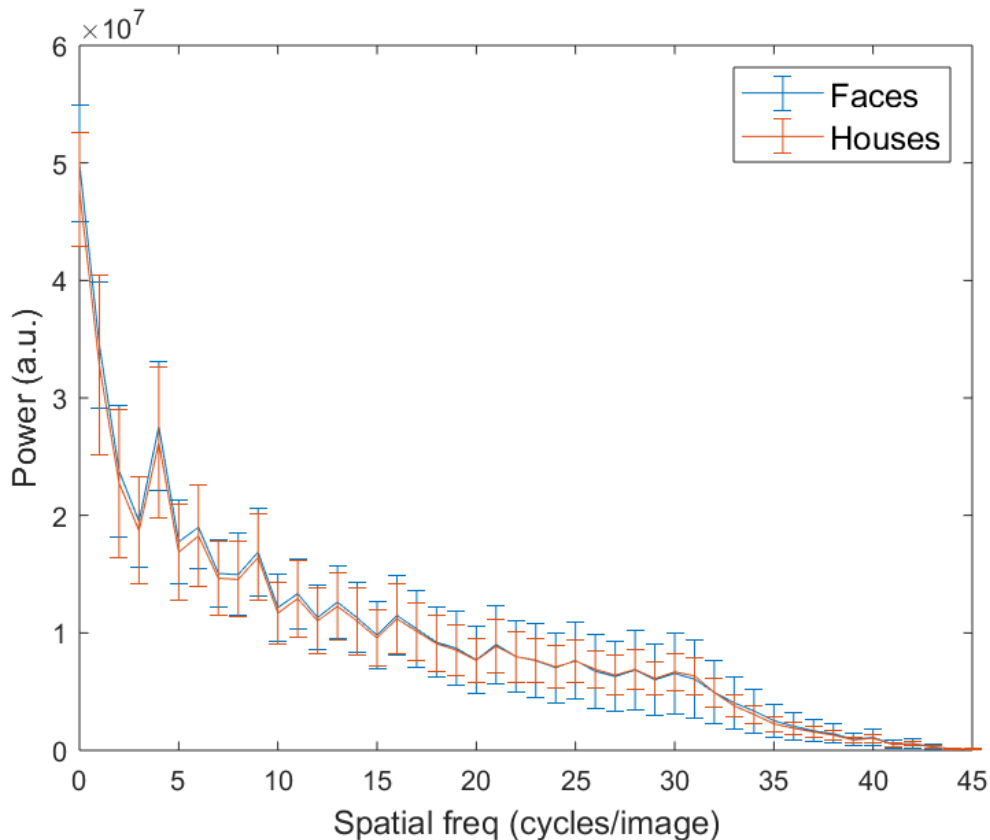


Figure 6. Power of the faces and houses at each spatial frequency (cycles/image).

2.5.2 Results

Reaction times for the fastest 10% of trials and the slowest 50% of trials can be found in Fig. 3D. As we obtained data from both target and foil trials, both were included in our analysis. In the fastest 10% of trials, faces were again detected more accurately than houses, despite the matching of spatial frequencies ($F(1,22)=4.83$, $p<0.05$) (Fig. 4D).

Again, the contribution of the retinocollicular pathway was assessed. No significant differences in accuracy were seen for the nasal compared with the temporal hemiretina ($F(1,22)=0.02$, $p=.890$). The interaction between hemiretina presentation and stimulus was also not significant; houses were not detected significantly more accurately than faces at fast reaction times when contrasting the temporal with the nasal hemiretina ($F(1,22)=.67$, $p=.421$) (Fig. 5D).

To investigate whether spatial frequency manipulation substantially modulated performance, the fastest 10% of trials from Experiment 3 were compared to the fastest 10% of trials from Experiment 4 using a 2x2 ANOVA with experiment and stimulus as the within-subject factors. Overall, there were no significant difference in performance between the two experiments ($F(1,22)=2.80, p=.108$). Furthermore, there was no significant interaction between stimulus and experiment, showing the difference in accuracy for faces compared with houses was not significantly different in Experiments 3 and 4 ($F(1,22)=0.60, p=.448$). In line with the results of each experiment, there was a main effect for stimuli, with faces detected significantly more accurately than houses at the fastest reaction times ($F(1,22)=10.17, p<0.01$).

To further investigate if response bias caused the differences in accuracy at the fastest reaction times, we calculated the false alarm and hit rate for the faces and houses. The mean false alarm rate was lower for faces ($M=.13, SE=0.018$) than for houses ($M=0.15, SE=0.018$) at the fastest 10% of RT's. The mean hit rate was higher for faces ($M=.91, SE=0.017$) than for houses ($M=.83, SE=0.017$) at the fastest 10% of RT's. A higher hit rate and a lower false alarm rate again confirms that response bias cannot account for the differences in accuracy, and that participants were better at identifying faces than houses.

2.5.3 Interim Discussion

This experiment explored the idea that spatial frequency might have caused the category-specific effects in Experiments 1 and 3. Altering the spatial frequency of the images did not have any significant effect on the results. In trials with fast RTs, faces were still detected more accurately than houses. Furthermore, the spatial filtering in Experiment 4 did not change the results from those seen in Experiment 3. These results support a face-specific rapid detection mechanism, rather than a low-spatial frequency mechanism.

2.6 General Discussion

Four experiments were conducted to determine whether a rapid route for face detection could be identified in a behavioural experiment. In addition we sought to determine whether a subcortical process, facilitated by the retinocollicular pathway, could be responsible for the rapid detection of faces. If the retinocollicular pathway to the

amygdala were responsible for rapid face detection, we would expect to see a benefit for faces, but not houses, when presenting to the nasal hemiretina. In Experiments 1, 3 and 4, participants rapidly detected faces but not houses from very distinct warped foil stimuli. However, there was no benefit of presenting the stimuli to the nasal hemiretina, providing no support for a retinocollicular route in rapid face detection in our task. Even when we combined the participants from Experiments 1 and 3 into a single analysis to increase power, we did not see a benefit for face detection in the nasal hemiretina, and in fact, faces or houses presented to the nasal hemiretina were detected less accurately.

We then considered what aspects of the face stimuli could have led to rapid detection. Faces have greater power at lower spatial frequencies than houses. In Experiment 4, we filtered the images to enhance relative power at high spatial frequencies for the faces and reduce it for the houses. Faces were still detected more accurately than houses, showing that it is category, and not just spatial frequency, that facilitates rapid detection. Furthermore, we found performance overall was no worse when high frequencies were emphasized. This suggests low spatial frequencies did not have a strong role, and that perhaps the rapid detection mechanism is capable of precise visual representation. We tested this in Experiment 2 and found that when participants were required to make fine visual discriminations, more accurate fast face detection disappeared. This suggests the rapid discrimination method is “dirty” as well as being “quick”. Again no contribution was evident from the retinocollicular pathway.

Taken together, our results show there is a rapid route for the detection of faces, which relies on coarse visual information, but not low on spatial frequencies in particular. In none of the experiments did we find evidence of a benefit for face detection in the nasal hemiretina. This could support the idea that a cortical rather than a subcortical mechanism is responsible for rapid face detection (Cauchoix, & Crouzet, 2013), and is congruent with evidence that cortex is capable of rapid processing (Barragan-jason et al., 2015; Foxe & Simpson, 2002; H. Liu et al., 2009). However, we acknowledge that no imaging (e.g. fMRI) was performed in this study. Therefore, although our behavioral experiment may motivate future imaging work, it does not provide the same evidence as imaging would to determine what brain structures are involved.

Strengthening the results of the study, the warped foil stimuli used in this experiment were well matched in terms of luminance, contrast, and spatial frequency to the target stimuli, and could not be differentiated in a model of the early visual system (HMAX, Stojanoski & Cusack, 2014), eliminating a series of confounding variables not often considered. There is one study where the authors found participants were orienting more quickly to ‘face like’ stimuli when they were presented to the nasal hemiretina (Tomalski, Johnson, & Csibra, 2009). However, this study used ‘Johnson faces’ where black boxes are put in place of the eyes, nose and mouth. Control stimuli in this study were an inverted version of the ‘Johnson face’. Although these control stimuli were matched for variables like spatial frequency, the target stimuli will have a large ‘top-heavy’ bias in comparison with the foil stimuli, which could be what was responsible for the increased performance of the nasal hemiretina. Our naturalistic stimuli will likely have had less of a top-heavy bias, and this could be a potential reason why we do not see a benefit of presenting to the nasal hemiretina. It is also possible that cortex is needed to make category judgements when target and foil stimuli are well matched.

In order to ensure that the visual stimuli were unrecognizable, Stojanoski and Cusack (2014) quantified how much warping was necessary to remove semantic information from different categories. Faces, along with bikes, needed the highest levels of warping in order to render them unrecognizable. Therefore, it is unlikely that face blocks in Experiments 1 and 3 represented an easier task than house blocks. In addition, different diffeomorphic fields were used for each foil, which makes them distinct, even within a category. For example, although the first exemplar in Figure 1 has a dark portion in the center, not all face stimuli have this. Across the entire face and house categories, these small characteristics were insufficient to have driven the broader differences between the faces and the houses.

Other researchers have also failed to see a benefit for face identification when presenting stimuli to the nasal hemiretina (Gabay, Burlingham, & Behrmann, 2014). In addition to the nasal/temporal manipulation, these researchers use a Wheatstone stereoscope to exploit the fact that visual information is segregated monocularly until visual cortex. Gabay et al. (2014) present stimuli monocularly, either to the same or different eyes and

have participants make identity judgements. They find a benefit for presenting stimuli to the same eye, which they hypothesize, could be due to the monocular properties of subcortical structures, such as the lateral geniculate nucleus (LGN). Although the retinocollicular pathway does not seem to be contributing to improved face detection, it is possible that the LGN, on the way to cortex, could be responsible for our results.

If the amygdala is not responsible to the rapid detection of faces, it could still be processing emotional information (Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). This could explain why we see decreased overall performance for the nasal hemiretina. It is possible that when the amygdala feeds information to the cortex creating competing processing, which makes it more difficult for cortex to rapidly categorize visual stimuli.

If faces are being detected more accurately at fast reaction times than the houses, what features of the stimuli are causing this increase in accuracy? Faces have significantly less inter-exemplar variability than houses. It is possible that the invariance of face stimuli allows tighter tuning in cortex, leading to more accurate, robust and efficient detection. If the invariability in our stimuli is causing the effects seen in the experiments, it is possible that other stimulus categories with limited variability could tap into a rapid mechanism. If other categories of stimuli could be capable of tapping into the rapid mechanism, are faces really special or is expertise what is important in order to develop “expert” face processing capabilities? Several studies have highlighted how important experience is in the processing of faces. For example, cataracts that substantially decrease visual input from reaching the right hemisphere in infancy impair “expert” face processing from completely developing (Le Grand, Mondloch, Maurer, & Brent, 2003). There is also evidence that perceptual narrowing and other complex aspects of face processing continue to emerge over the first year of life, substantiating the hypothesis that experience is important in face processing (Kelley, Quinn, Slater, Lee, Ge, Pascalis, 2008; Sai, 2005). However, other researchers have found that cortex responds to faces extremely quickly after birth (Tzourio-Mazoyer et al., 2002), and dispute the experience hypothesis (McKone, Crookes, Jeffery, & Dilks, 2012). Our results could suggest that

other categories of stimuli with limited variability and increased experience could tap into this rapid route.

Another theoretical framework that our results could be considered in is dual process theory. Proponents of dual process theories have suggested that there are two processes involved in cognition, the first an unconscious process (often thought of as procedural learning), and the second a conscious, effortful process (i.e. explicit learning) (Barrett, Tugade, & Engle, 2004). In the categorization literature, others have proposed a dual process model specific to categorization (COVIS), which has a procedural learning component and a cognitively demanding, verbal hypothesis driven component, mediated by the executive network (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Maddox & Ing, 2005). It could be that faces access a rapid procedural mechanism while slower categorization is dominated by the explicit process. However, other researchers have criticized COVIS (Newell, Dunn, & Kalish, 2011). From our current data is impossible to determine whether our results are reflective of a dual process theory or are the result of a single process that is more robust to faces. Future work should seek to examine this.

2.7 Conclusions

In conclusion, faces were detected with greater accuracy at fast reaction times than houses, when they are distinct from the foil stimuli. Our data do not offer any support that these results are due to the contributions of the retinocollicular pathway, suggesting that an alternative route to cortex is involved in the rapid detection of faces.

2.8 References

- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–81. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9697427>
- Awasthi, B., Sowman, P. F., Friedman, J., & Williams, M. A. (2013). Distinct spatial scale sensitivities for early categorization of faces and places: neuromagnetic and behavioral findings. *Frontiers in Human Neuroscience*, *7*(March), 1–11. <http://doi.org/10.3389/fnhum.2013.00091>
- Barragan-jason, G., Cauchoix, M., & Barbeau, E. J. (2015). The neural speed of familiar face recognition. *Neuropsychologia*, *75*, 390–401. <http://doi.org/10.1016/j.neuropsychologia.2015.06.017>
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, *130*(4), 553–573. <http://doi.org/10.1037/0033-2909.130.4.553>
- Benevento, L. A., & Standage, G. P. (1983). The organization of projections of the retinorecipient and nonretinorecipient nuclei of the pretectal complex and layers of the superior colliculus to the lateral pulvinar and medial pulvinar in the macaque monkey. *The Journal of Comparative Neurology*, *217*(3), 307–336. <http://doi.org/10.1002/cne.902170307>
- Bertini, C., Leo, F., & Làdavias, E. (2008). Temporo-nasal asymmetry in multisensory integration mediated by the Superior Colliculus. *Brain Research*, *1242*(MAY), 37–44. <http://doi.org/10.1016/j.brainres.2008.03.087>
- Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, *26*(1), 157–170.
- Cauchoix, S Crouzet, M. C. (2013). How plausible is a subcortical account of rapid visual recognition? *Frontiers in Human Neuroscience*, *7*(February), 1–4. <http://doi.org/10.3389/fnhum.2013.00039>
- Celebrini, S., Thorpe, S., Trotter, Y., & Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake primate. *Visual Neuroscience*, *10*, 811–825.
- Conley, M., Lachica, E. A., & Casagrande, V. A. (1985). *Demonstration of ipsilateral*

- retinocollicular projections in the tree shrew (Tupaia glis)*. *Brain Research* (Vol. 346).
- Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision*, *10*(2010), 16.1-17.
<http://doi.org/10.1167/10.4.16>
- Delorme, a., & Thorpe, S. J. (2001). Face identification using one spike per neuron: Resistance to image degradations. *Neural Networks*, *14*(6–7), 795–803.
[http://doi.org/10.1016/S0893-6080\(01\)00049-1](http://doi.org/10.1016/S0893-6080(01)00049-1)
- Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*, 571–584.
<http://doi.org/10.1068/p5705>
- Foundation, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). <http://doi.org/10.1126/science.aac4716>
- Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans A framework for defining “early” visual processing, 139–150.
<http://doi.org/10.1007/s00221-001-0906-7>
- Gabay, S., Burlingham, C., & Behrmann, M. (2014). The nature of face representations in subcortical regions. *Neuropsychologia*, *59C*, 35–46.
<http://doi.org/10.1016/j.neuropsychologia.2014.04.010>
- Garvert, M. M., Friston, K. J., Dolan, R. J., & Garrido, M. I. (2014). Subcortical amygdala pathways enable rapid face processing. *NeuroImage*, *102*, 309–316.
<http://doi.org/10.1016/j.neuroimage.2014.07.047>
- Harrison, D. W. (2015). *Brain Asymmetry and Neural Systems*. Cham: Springer International Publishing. <http://doi.org/10.1007/978-3-319-13069-9>
- Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10827445>
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews. Neuroscience*, *6*(10), 766–74. <http://doi.org/10.1038/nrn1766>
- Jolij, J., & Lamme, V. A. F. (2005). Repression of unconscious information by conscious processing: Evidence from affective blindsight induced by transcranial magnetic

- stimulation. *Proceedings of the National Academy of Sciences*, *102*(30), 10747–10751. <http://doi.org/10.1073/pnas.0500834102>
- Jones, E. G., & Burton, H. (1976). *A projection from the medial pulvinar to the amygdala in primates. Brain Research* (Vol. 104).
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area : A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area : a cortical region specialized for the perception of faces, (November), 2109–2128. <http://doi.org/10.1098/rstb.2006.1934>
- Kelley, D, Quinn, P, Slater, A, Lee, K, Ge, L, Pascalis, O. (2008). The Other-Race Effect Develops During Infancy, *18*(12), 1084–1089. <http://doi.org/10.1111/j.1467-9280.2007.02029.x>.The
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. <http://doi.org/10.1016/j.visres.2005.10.002>
- Konig, P., Engel, A. K., Roelfsema, P. R., & Singer, W. (n.d.). How Precise is Neuronal Synchronization?
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience*, *6*(10), 1108–12. <http://doi.org/10.1038/nn1121>
- Leopold, D. a, & Rhodes, G. (2010). NIH Public Access, *124*(3), 233–251. <http://doi.org/10.1037/a0019460.A>
- Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Article Timing , Timing , Timing : Fast Decoding of Object Information from Intracranial Field Potentials in Human Visual Cortex. *Neuron*, *62*(2), 281–290. <http://doi.org/10.1016/j.neuron.2009.02.025>
- Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience*, *5*(9), 910–916. <http://doi.org/10.1038/nn909>
- Luo, Q., Holroyd, T., Majestic, C., Cheng, X., Schechter, J., & Blair, R. J. (2010). Emotional Automaticity Is a Matter of Timing, *30*(17), 5825–5829.

<http://doi.org/10.1523/JNEUROSCI.BC-5668-09.2010>

- Maddox, T. W., & Ing, D. A. (2005). Delayed Feedback Disrupts the Procedural-Learning System but Not the Hypothesis-Testing System in Perceptual Category Learning. *Journal of Experimental Psychology*, *21*(1), 100–107.
<http://doi.org/10.1037/0278-7393.31.1.100>
- Martin, C. B., McLean, D. A., O’Neil, E. B., & Köhler, S. (2013). Distinct familiarity-based response patterns for faces and buildings in perirhinal and parahippocampal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(26), 10915–23. <http://doi.org/10.1523/JNEUROSCI.0126-13.2013>
- McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: experience is less important than previously believed. *Cognitive Neuropsychology*, *29*(1–2), 174–212.
<http://doi.org/10.1080/02643294.2012.660138>
- Mende-Siedlecki, P., & Verosky, S. (2013). Robust selectivity for faces in the human amygdala in the absence of expressions, 2086–2106. <http://doi.org/10.1162/jocn>
- Méndez-Bértolo, C., Moratti, S., Toledano, R., Lopez-Sosa, F., Martínez-Alvarez, R., Mah, Y. H., ... Strange, B. A. (2016). A fast pathway for fear in human amygdala. *Nature Neuroscience*, *19*(8), 1041–1049. <http://doi.org/10.1038/nn.4324>
- Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of Category Learning: Fact or Fantasy? In *The Psychology of Learning and Motivation* (Vol. 54, pp. 167–215). Academic Press. <http://doi.org/10.1016/B978-0-12-385527-5.00006-1>
- Perry, V. H., & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey’s retina: implications for central magnification factors. *Vision Research*, *25*(12), 1795–810. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3832605>
- Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*, *56*(4), 2356–2363.
<http://doi.org/10.1016/j.neuroimage.2011.03.067>
- Pitcher, D., Walsh, V., Yovel, G., & Aviv, T. (2007). TMS Evidence for the Involvement of the Right Occipital Face Area in Early Face Processing. *Current Biology*, *17*, 1568–1573. <http://doi.org/10.1016/j.cub.2007.07.063>

- Pollack, J. G., & Hickey, T. L. (1979). The distribution of retino-collicular axon terminals in rhesus monkey. *The Journal of Comparative Neurology*, *185*(4), 587–602. <http://doi.org/10.1002/cne.901850402>
- Rafal, R. D., Koller, K., Bultitude, J. H., Mullins, P., Ward, R., Mitchell, A. S., & Bell, A. H. (2015). Connectivity between the superior colliculus and the amygdala in humans and macaque monkeys: virtual dissection with probabilistic DTI tractography. *Journal of Neurophysiology*, *114*(3), 1947–62. <http://doi.org/10.1152/jn.01016.2014>
- Sadeh, B., Podlipsky, I., Zhdanov, A., & Yovel, G. (2010). Event-related potential and functional MRI measures of face-selectivity are highly correlated: A simultaneous ERP-fMRI investigation. *Human Brain Mapping*, *31*(10), 1490–1501. <http://doi.org/10.1002/hbm.20952>
- Sai, F. Z. (2005). The Role of the Mother 's Voice in Developing Mother 's Face Preference : Evidence for Intermodal Perception at Birth, *50*, 29–50. <http://doi.org/10.1002/icd>.
- Sato, W., Kochiyama, T., Uono, S., Matsuda, K., Usui, K., Inoue, Y., & Toichi, M. (2013). Rapid and multiple-stage activation of the human amygdala for processing facial signals. *Communicative & Integrative Biology*, *6*(4), e24562. <http://doi.org/10.4161/cib.24562>
- Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, *104*.
- Sterling, P. (1973). Quantitative mapping with the electron microscope: retinal terminals in the superior colliculus. *Brain Research*, *54*, 347–354.
- Stojanoski, Bobby, Cusack, R. (2014). Time to wave good-bye to phase scrambling : Creating controlled scrambled images using diffeomorphic transformations, *14*, 1–16. <http://doi.org/10.1167/14.12.6.doi>
- Sylvester, R., Josephs, O., Driver, J., & Rees, G. (2007). Visual fMRI Responses in Human Superior Colliculus Show a Temporal – Nasal Asymmetry That Is Absent in Lateral Geniculate and Visual Cortex, (Fries 1984), 1495–1502. <http://doi.org/10.1152/jn.00835.2006>.
- Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of

- emotional signals. *Nature Reviews. Neuroscience*, *11*(10), 697–709.
<http://doi.org/10.1038/nrn2889>
- Tamietto, M., Pullens, P., De Gelder, B., Weiskrantz, L., & Goebel, R. (2012).
 Subcortical connections to human amygdala and changes following destruction of
 the visual cortex. *Current Biology*, *22*, 1449–1455.
<http://doi.org/10.1016/j.cub.2012.06.006>
- Tomalski, P., Johnson, M. H., & Csibra, G. (2009). Temporal-nasal asymmetry of rapid
 orienting to face-like stimuli. *Neuroreport*, *20*(15), 1309–1312.
<http://doi.org/10.1097/WNR.0b013e32832f0acd>
- Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000).
 Response properties of the human fusiform face area. *Cognitive Neuropsychology*,
17(1), 257–80. <http://doi.org/10.1080/026432900380607>
- Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer,
 B. (2002). Neural correlates of woman face processing by 2-month-old infants.
NeuroImage, *15*(2), 454–61. <http://doi.org/10.1006/nimg.2001.0979>
- VanRullen, R., Guyonneau, R., & Thorpe, S. J. (2005). Spike times make sense. *Trends
 in Neurosciences*, *28*(1), 1–4. <http://doi.org/10.1016/j.tins.2004.10.010>
- Vanrullen, R., & Koch, C. (2001). Visual Selective Behavior Can Be Triggered by a
 Feed-Forward Process. *Journal of Cognitive Neuroscience*, *15*(2), 209–217.
- Visconti di Oleggio Castello, M., & Gobbini, M. I. (2015). Familiar Face Detection in
 180 ms. *PloS One*, *10*(8), e0136548. <http://doi.org/10.1371/journal.pone.0136548>
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial
 frequency sensitivities for processing faces and emotional expressions. *Nature
 Neuroscience*, *6*(6), 624–31. <http://doi.org/10.1038/nn1057>
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004).
 Distant influences of amygdala lesion on visual cortical activation during emotional
 face processing, *7*(11), 1271–1278. <http://doi.org/10.1038/nn1341>

Chapter 3

3 Category-selective visual regions have a distinct signature of connectivity early in infancy.

By four months, infants can form categories of similar looking objects, but it is unclear when they begin to make the rich cross-modal, motoric and affective associations that are characteristic of adult visual categories. These associations are thought to be encoded by long-range brain connectivity, and is reflected in the distinctive signature of connectivity of each category-selective region in the ventral visual stream. Category-selective ventral visual regions are already functioning in young infants, but their long-range connectivity has not been investigated. Therefore we used MRI diffusion tractography to characterize the connectivity of face, place and tool regions in 1-9 month infants. Using a linear discriminant classifier, we found that the face and place regions had adult-like connectivity throughout infancy, but the tool-network underwent significant maturation until 9 months. This suggests that young infants might already be forming rich associations, but that different categories are maturing with different developmental trajectories.

3.1 Introduction

Forming categories is a core part of human development as it allows us to improve our inferences about the environment and make better predictions. At 3-4 months old, infants are able to form visual, perceptual categories (e.g. cats vs. dogs) (Eimas & Quinn, 1994; Quinn, Eimas, & Rosenkrantz, 1993). If they are presented with a series of cats, they will preferentially orient towards a subsequently presented dog, which shows that they can form a stable representation of the distribution of perceptual features of the cat stimuli (Sloutsky & Fisher, 2004; Wilcox, 1999) and detect deviations from the category. The visual processing that leads to these perceptual categories is thought to be located in the ventral occipital and temporal cortex. In adults, distinct regions are selective for perceptual categories such as faces, places or objects (Epstein, 2008; Epstein & Kanwisher, 1998; Nancy Kanwisher & Yovel, 2006; N. Kanwisher, Woods, Iacoboni, &

Mazziotta, 1997; Malach et al., 1995). Functional neuroimaging has found that at 4-6 months infants already have regions that are selective for perceptual categories in the ventral temporal cortex as well (Deen et al., 2017).

In adults, the pattern of connectivity to other brain areas is distinct for each category-selective region in the ventral temporal cortex (Osher et al., 2016; Saygin & Kanwisher, 2014). These long-range connections are thought to encode the cross-modal, motoric and affective associations characteristic of rich semantic categories (Binder, Desai, Graves, & Conant, 2009; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Patterson, Nestor, & Rogers, 2007). As concrete examples, seeing a silent video of a dog barking evokes the representation of its sound in auditory cortex (Meyer, Kaplan, Essex, Webber, Damasio & Damasio, 2010), and for tools and objects, category representations in the ventral stream are integrated with action representations in the dorsal stream (Almeida, Fintzi, & Mahon, 2013; Coccia, Bartolini, Luzzi, Provinciali, & Ralph, 2004; Goodale & Milner, 1992). The importance of long-range connections in semantics was recently demonstrated by multivariate decoding of white matter pathways in brain-injured patients with semantic deficits (Fang et al., 2018).

It has not been established when infants begin go beyond visual, perceptual categories to make rich associations. It is difficult to measure, as conceptual understanding is usually probed verbally, but infants in the first year have very limited language. Could they already, however, be forming associations through a process of unsupervised statistical learning? In this study we use neuroimaging to investigate this question, by examining the maturity of structural connectivity of category-selective regions in infants. To measure connectivity, we used diffusion-weighted imaging and tractography. We extracted the characteristic signatures of connectivity of three category-selective regions in adults using a machine learning approach, and then tested for generalisation to infants.

3.2 Results

In order to measure structural connectivity, diffusion-weighted images were obtained from 14 adults and 11 infants. Probabilistic tractography was performed using seed and target regions taken from the parcellation by the Human Connectome Project (HCP)

(Glasser et al., 2016). Each voxel in the ventral visual stream, as defined by the HCP (see Methods), was used as a seed, while the brain areas outside the ventral stream acted as targets for tractography. To define category-selective regions in the ventral stream, contrast maps from the HCP fMRI localizers were used to determine the regions within the parcellation that were most selective for face, places and tools. These regions were the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3, respectively. Regions of interest are displayed in Figure 7a.

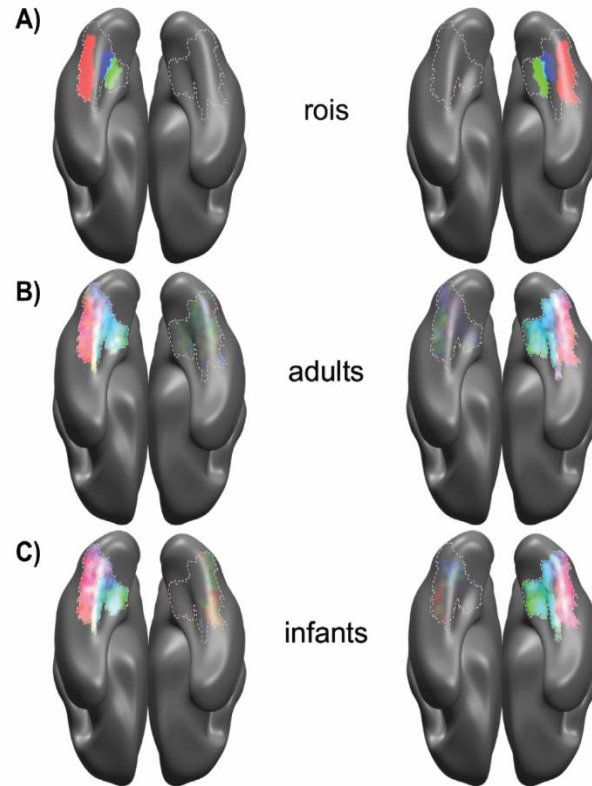


Figure 7. Regions used for classification and group overlays of the voxels selected by the classifiers as part of the face, place, and tool regions of interest based on their structural connectivity with other brain regions.

A) Regions from the HCP parcellation that were most selective for faces, tools and places (red, blue and green, respectively) in the left and right hemispheres (left and right columns). Dotted outlines represent the ventral stream seed region, as defined by the HCP (see methods) B) Voxels identified by a linear-discriminant classifier as selective for faces, places, and tools in adult participants (N=14), based on their distinctive signature of structural connectivity with the rest of the brain. Classification was performed separately for the left and the right hemisphere, using leave-one-subject-out cross validation. Group average overlay maps are shown with the same color mapping as (A). C) The distinctive signatures of structural connectivity were also present in infants (N=11), as shown by voxels identified as category-selective by a linear-discriminant classifier trained on adult connectivity and tested in infants.

The connectivity pattern for the category-selective regions as then probed using three linear-discriminant classifiers, one for each visual category. Using leave-one-subject-out cross-validation, a classifier was trained to differentiate voxels from the category selective regions from the other voxels in the ventral stream, based on their structural connectivity with the rest of the brain. The classifier's performance was then tested on the left out subject. Using signal detection theory, d-primes were calculated for each participant, to determine how sensitive the classifiers were in locating voxels in the face, place and tool regions. All three regions could be robustly localized in adults ($t(13)=26.26, p<0.001, t(13)=22.35, p<0.001, t(13)=17.17, p<0.001$) (Figure 7b). Classification performance is quantified in Figure 8a, which shows the d-primes for classification of the imaging data.

In order to characterize the connectivity of the infant ventral stream, probabilistic tractography was also performed on the infant diffusion data, using the same seed and target regions as in adults (see Methods for details of two-stage adult-to-infant normalization procedure). To determine whether the category-selective regions that were present in adults were present in infants, linear discriminant classifiers were trained on the entire adult dataset in the manner described above. These classifiers were then tested on the infant data. The classifiers localized all three regions in the infants ($t(10)=24.47, p<0.001, t(10)=16.54, p<0.001, t(10)=5.95, p<0.001$) (Figure 7c). However, there was a category-by-group interaction ($F(2,46)=6.64, p<0.01$). Post-hoc tests showed this was because the face and place regions were as strongly detected in infants as they were in adults ($t(23)=0.165, N.S., t(23)=0.257, N.S.$), but the tool region was detected with greater accuracy in adults than in infants ($t(23)=3.62, p<0.01$) (Figure 8b). Finally, in order to examine the developmental trajectory of the networks, infant age and classification accuracy (d-primes) were correlated-- only the tool network underwent significant change over the first 9 months of postnatal life (faces: $r(9)=-0.03, N.S.$; tools: $r(9)=0.75, p<0.01$; places: $r(9)=-0.01, N.S.$) (Figure 9).

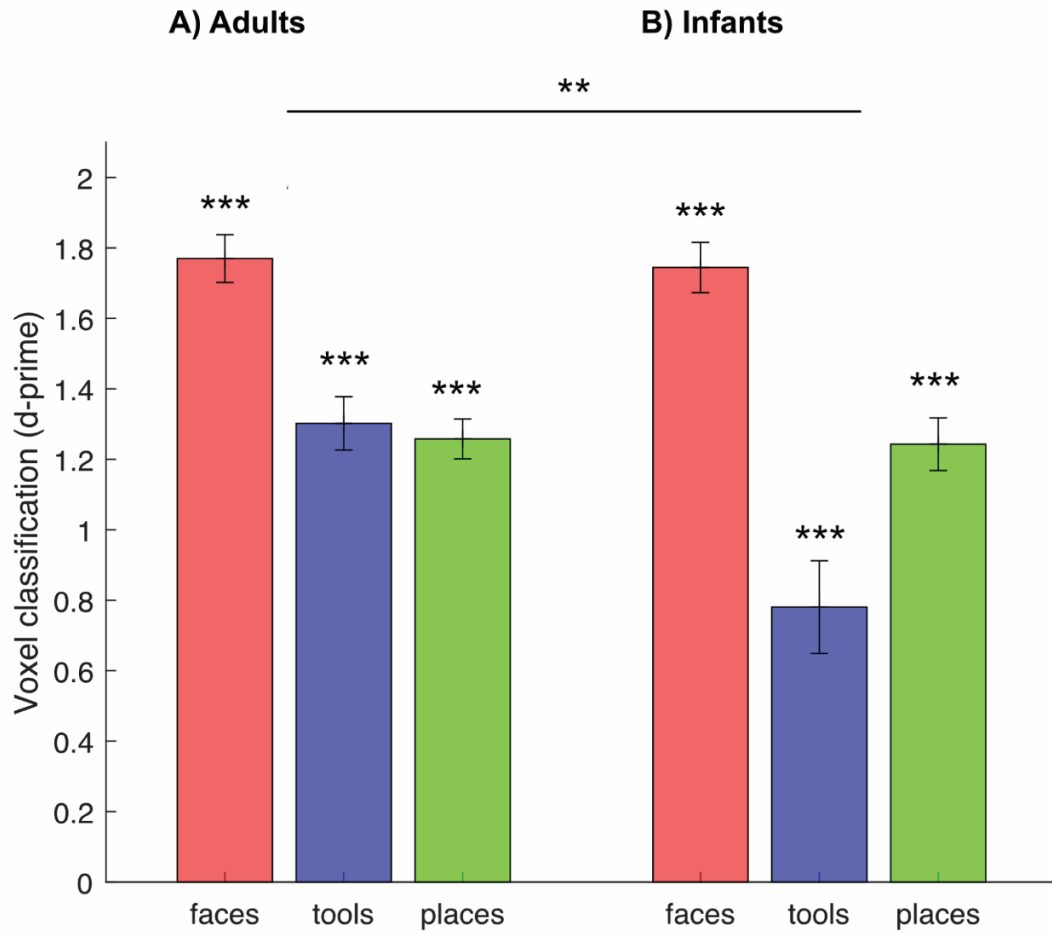


Figure 8. Voxel classification performance for infants and adults

a) Voxel classification performance for the adults (N=14) quantified using d-prime, collapsed across hemispheres. b) Voxel classification performance for the infants (N=11) again measured using d-prime. All regions were robustly localized in infants and adults but there was a significant difference in detection accuracy between the infant and adult tool region, demonstrating the immaturity of connectivity for the tool region during infancy. The mean +/- one standard error across subjects is shown.

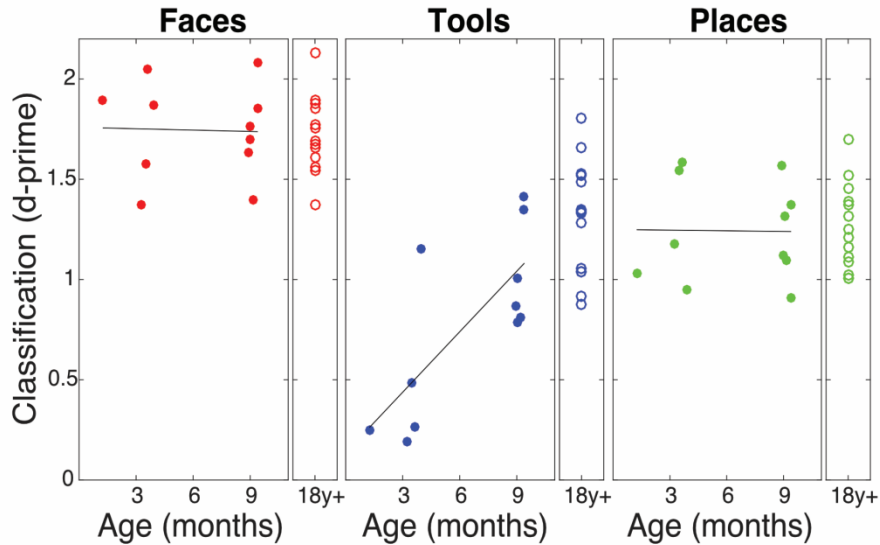


Figure 9. The relationship between the age of participants (14 adults and 11 infants) and classification accuracy (d-prime) with best-fit lines.

Only tool classification had a significant relationship with age, demonstrating the maturation of the distinctive connectivity of the tool network over the first year of postnatal life

These results suggest the connectivity of the tool region develops later than that of the face and place regions, but we also examined an alternative explanation. Could it be that tractography is more difficult in infants than adults, because of their lower signal-to-noise or smaller brains, and that identification of voxels in the tool region is more sensitive to this? Two analyses were conducted to investigate this. First, in adults, the detection of the tool-selective voxels was no worse than detection of the place-selective voxels and performance was not at ceiling (Figure 8A), suggesting that detection of the tool region is not intrinsically more difficult. Second, we compared region size, which may affect performance more strongly in smaller infant brains; the place and tool regions were the same size in one hemisphere and were less than 10 voxels different in the other. The hit rate for the place and tool region was also not significantly different in adults ($t(13) = -1.04, N.S.$).

We next examined which target regions of connectivity most strongly influenced each of the classifiers (see Appendix A). For the place area, this was connectivity to a network strongly associated with navigation, including the hippocampus, parahippocampal areas,

and the entorhinal cortex (Epstein, 2008; Hafting, Fyhn, Molden, Moser, & Moser, 2005). The face region's strongest-weighted structural connections were to area PH, which is strongly deactivated in the HCP functional face contrasts (Glasser et al., 2017). The tool region showed strong weighting of connections to visual regions, and accordingly, tools often have distinctive basic perceptual features (Tyler & Moss, 2001). Following this tools showed strong connections to the 4th visual area, which is associated with color processing. Tools also showed strong connections to the third visual area, which is connected to the posterior parietal regions that are associated with visuomotor transformations. Finally, the tool region's strong connectivity with the posterior orbitofrontal complex (OFC) may be driven by top-down tool classification (Bar et al., 2006).

The HCP tool region was located in the cortex between the place and face regions. Although tool selectivity has been found before in this location (Chao, Haxby, & Martin, 1999) it is also present in other areas (Grill-Spector & Weiner, 2014). From Figure 1b, it is apparent that even in the adults, there is some blurring between category boundaries, particularly between the tool and place regions. As the three classifiers were set up to each independently discriminate a single category selective region from all other voxels (including those that were selective for no category), these results cannot be used to quantify if pairs of categories can be distinguished from each other. To address this, we repeated the classification, but with a fitted discriminant analysis classifier that allowed for multiclass classification (Guo, Hastie, & Tibshirani, 2007). Using multiclass classification meant a single classifier aimed to predict whether a voxel was face, place, tool or non-category selective. This confirmed that the three category-selective regions could be robustly discriminated from each other with even the smallest pairwise difference in d' , for tools vs. places, reliable in adults ($t(13)=7.05, p<0.001$) and infants ($t(10)=2.40, p<0.05$).

3.3 Discussion

All three category-selective regions could be robustly localized in adults. These results extend the fMRI results found by Deen et al. (2017) and the structural connectivity patterns found in adults by Saygin and Kanwisher (2014), demonstrating adult ventral

stream category-selective regions can be delimited based on their distinctive signature of structural connectivity with the rest of the brain. Additionally, the broader networks associated with the category-selective regions were found to be completely or partially mature in infants, with the tool network experiencing a longer maturational time course, extended over the first 9 months of postnatal life. The early maturity of the distinctive networks associated with category-selective regions suggests that infants might be going beyond perceptual categories before they are able to express themselves verbally. The associations of the three perceptual categories are reflected in the connections identified by the classifiers. For example, many of the strongest connections for the place region are part of the place processing network—the hippocampus, entorhinal cortex and parahippocampal regions (see S1) (Epstein, 2008). However, we do acknowledge that mature connectivity profiles do not necessitate that infants have fully formed semantic, multimodal or conceptual representations. In fact, it's likely that they do not, and that networks continue to mature well into the synaptic pruning that occurs during adolescence.

It is likely that experience is playing a substantial role in the transition from perceptually based categories to conceptual maturation. For example, by one month, infants have had considerable experience with faces, spending an estimated quarter of their waking hours with faces taking up the majority of their visual field (Jayaraman, Fausey, & Smith, 2015). In contrast with this experience-driven maturation hypothesis, other researchers have proposed an innate face processing module in the brain. The most recent support for innate face processing comes from (Reid et al., 2017) who demonstrated that fetuses will preferentially orient to face-like patterns in the third trimester. However, these results have generated controversy (Scheel, Ritchie, Brown, & Jacques, 2017) and does not align with a study in infant monkeys, which found that experience with faces is necessary to develop a typical adult like face processing system (Arcaro, Schade, Vincent, Ponce, & Livingstone, 2017).

Infants also have substantial experience viewing scenes (and due to their long supine hours, perhaps particularly ceilings). The scene representations in the ventral visual stream are biased towards the periphery, and the retinal temporal hemifield, representing

the periphery, develops before the nasal hemifield (representing more foveal representations) (Grill-Spector & Weiner, 2014; Lewis & Maurer, 1992). This combination of experience and mature brain circuitry might drive the mature place network connectivity, and therefore possible conceptual maturation, seen in the study.

In contrast, young infants are perhaps less likely to see tools than faces and places. Tools also are a less homogenous category, which will make category-level recognition more difficult, especially in comparison with faces, which have high similarity between exemplars. Compared with categories that are passively perceived, tool use also requires integration between sensory and motor representations, which might require more extensive experience with the environment. By 9 months, infants are able to differentiate between textures (Bushnell & Boudreau, 1993), and hold spoons correctly during self-feeding (Bushnell & Boudreau, 1993). Using head mounted eye trackers, researchers have determined that once infants learn to reach, they'll often hold an object quite close to their face, which has been shown to be an ideal training stimulus for neural networks to recognize objects (Bambach, Crandall, Smith, & Yu, 2018). In line with this evidence and the prolonged maturation of the tool network observed here, researchers have proposed a perception-action theory, where interactions between perception and motor experience gradually accrued over development explain the maturation of tool use (Kahrs & Lockman, 2014; Lockman, 2000). This theory is supported by behavioral evidence that experience with tools drives tool use behavior (Barrett, Davis, & Needham, 2007).

The new methods developed in this work were used for the first time in infant neuroimaging. These methods have the potential as a diagnostic method for infants at high risk of developing impairments due to neurological abnormalities. In adults and children, connectivity has been shown to be predictive of many brain disorders, including mental health disorders, and can be more predictive than the standard measures doctors will use to prescribe treatment (Fox, 2018; Whitfield-Gabrieli et al., 2016). Early diagnosis could facilitate early intervention, when the brain is most plastic, which may have promising potential for infants, their families and society as a whole.

A limitation of our results is the lack of functional and diffusion data in the same participants. Future research may be able to more closely identify the network maturation of category-selective regions with a comprehensive longitudinal study, where functional localizers are acquired in both awake infant and adult participants. However, one advantage to using the regions derived from the HCP is their generalizability across a large group of participants, something that would be challenging to do in a local sample. The HCP region definitions are also based on multiple types of data (structural, functional and diffusion data), which would also be challenging to acquire locally in large numbers in infants and adults. Nevertheless, these efforts would be worthwhile and could help answer many outstanding questions.

3.4 Methods

3.4.1 Data Acquisition

For both the adult and infant participants, high-quality diffusion-weighted MRI data were acquired using a 3T Siemens Prisma Magnetron Scanner at the Centre for Functional and Metabolic Mapping of Western University. Using a 20-channel head coil, the Minnesota multiband sequence was used (128 directions, 2 mm isotropic, no gap between slices, $b=1500 \text{ mm}^2 \text{ s}^{-2}$, multiband acceleration 4, monopolar diffusion encoding gradients, time of acquisition: 9 min and 18 sec). Using monopolar diffusion encoding gradients creates larger eddy currents, which distort the magnetic field and cause image distortion. Using the solution developed by the HCP, two scans were acquired with opposite phase-encoding polarities (left-to-right and one right-to-left). Combining these images during analysis with FSL's TOPUP calculates the susceptibility distortion, while EDDY corrects for eddy current-induced distortions and participant movement.

During the scan, younger infants were swaddled and wrapped in a Medvac pillow bag to help them remain still. Infants older than 6 months were not swaddled. All infants wore Mini Muffs adhesive sound protection (Natus, 7dB attenuation) and ear defenders (29 dB attenuation). Infants were scanned during natural sleep. Adult participants wore standard ear plugs and ear defenders and were requested to be as still as possible.

3.4.2 Participants

Diffusion-weighted MRI acquisitions were available from 11 sleeping infants as part of a larger infant imaging project with 51 participants. Infants were recruited either through public advertising or through clinical collaborators at the neonatal intensive care unit in London, Ontario. Diffusion MRI was acquired in 14 infants but three were subsequently excluded because of apparent brain injury. This left six healthy controls and four low-risk infants born preterm. One infant was scanned twice, but as the scans were two months apart, they were treated as separate participants in the analysis, making for a total of 11 infant datasets. Clinical information for the premature infants was obtained from medical records and a radiologist reviewed each scan for suspected brain injury. Infants were between 1 and 9 months old (corrected-age for infants born preterm, $M=6.4$ months, $SD=3.2$ months).

Diffusion-weighted MRI was also acquired from 16 adults at Western University. Participants were between 18 and 40 years old ($M=22.75$, $SD=4.89$). Author LC participated in the study and her data is included in the analysis. One participant was excluded because of an incidental finding, while another was excluded due to technical difficulties.

Approval for the study was provided by the Western University's Health Sciences Research Ethics Board. All parents provided informed consent before infants were scanned. All adult participants also provided informed consent.

3.4.3 Preprocessing

The data was analysed with a pipeline built from the automatic analysis (*aa*) software, FSL, and custom Matlab (R2016a). *aa* divides the description of the analysis into a user script that describes what data should be analysed, the study specific settings, and a task-list. This user script then calls the *aa* engine, which runs the processing pipeline, ensuring that only the stages not already completed are run, and that when possible modules are executed in parallel. The task list describes which processing modules should be used to analyse the data. For this analysis, the modules identified the DICOM files and organized them based on header information (*aamod_autoidentifyseries_timtrio*,

aamod_get_dicom_diffusion) and converted them to NIFTI format (*aamod_convert_diffusion_phaseencode_direction*). Then, *aamod_diffusion_extractnodif* identified the 10 volumes where $b=0$ in the diffusion data. The following six stages called components of the FSL diffusion processing pipeline. To combine the negative and positive phase encoding diffusion data into a single image and reduce distortion, *aamod_diffusion_topup* (TOPUP). *aamod_bet_diffusion* then removed non-brain tissue in the $b=0$ image (BET). In order to correct for any residual distortions due to eddy currents or head motion, *aamod_diffusion_eddy* (EDDY). *aamod_diffusion_dtifit* was then run to model diffusion tensors at the voxel level (DTIFIT).

In order to obtain mappings between individual brains to standard (MNI) space for the infants and adults, the normalization procedure for FSL's tract-based spatial-statistics (TBSS) was run. This normalizes the fractional anisotropy (FA) image to a mean FA tract skeleton, using non-linear registration. Normalizing the FA image resulted in a good registration for both the infant and adult data. Conventional normalization using a structural (T1 or T2) image to a template was not possible for the infants, as a number of the structural images were of poor quality due to participant motion. To ensure that normalizing with FA was not introducing an artifact into our results, additional analyses (not shown) in the adults where good structural images were also available, confirmed that very similar results were obtained if normalization was performed using the structural rather than diffusion images.

3.4.4 Human Connectome Project

In order to identify seed and target regions for the diffusion analysis, the parcellation from the HCP was used. The HCP parcellation segments the brain into 180 distinct regions in each hemisphere, based on structural, functional and diffusion data. To identify the seed regions, the HCP definition of the ventral stream visual cortex (region 4, supplementary neuroanatomical results) was used to identify the 14 regions that make up the ventral visual stream (Glasser et al., 2016). The individual voxels that were part of the 14 regions in the ventral visual stream were used as seeds and were excluded from tractography targets. The other 346 regions from the parcellation served as the target regions in the analysis. These seed and target regions were projected from the cortical

surface into volumetric MNI space. The normalization parameters from TBSS were then used to project these regions from MNI space into each subject's individual diffusion data space for tractography.

To select the regions in the ventral visual stream which were most responsive to faces, places, and tools, the functional MRI localizers from the HCP project were used. The category-average contrasts were used to select regions. These regions were the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3, respectively. Regions can be seen in Figure 7A. These regions were used as the category-selective regions in the subsequent classification analysis.

3.4.5 Tractography and Classification

Using the data from aa's *aamod_diffusion_bedpostx* module (BEDPOSTX), probabilistic tractography was performed in the infants and adults using FSL's PROBTRACKX, using 5,000 streamlines per seed voxel in the individual subject's ventral visual stream. The output of PROBTRACKX was then transformed to MNI space. These results were then summarized into a connectivity matrix that contained, for each voxel in the MNI ventral visual stream seed region the number of streamlines that terminated in each of the 346 target ROIs.

Three linear discriminant classifiers were then trained to identify the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3 in adults, based on connectivity with the 346 target regions. For the adults, leave-one-subject out cross validation was used to test whether selectivity could be predicted from connectivity. For the infants, three classifiers trained in a similar way on the full adult dataset were then tested on the infant diffusion data. D-primes were calculated to evaluate the accuracy of the classifiers in both infants and adults. To test for a relationship between age and connectivity in the infants, Pearson correlations were calculated between age and d-prime scores for each category.

For the multiclass classification, a fitted discriminant analysis classifier (www.mathworks.com/help/stats/fitdiscr.html) was used to identify the fusiform

complex, the ventromedial visual area 2, and ventromedial visual area 3, as well as the non-category selective voxels in the HCP ventral stream visual cortex (region 4, supplementary neuroanatomical results) in adults. This was done based on voxel-wise connectivity with the 346 target regions. Leave one-out cross validation was used to test classification accuracy, and d-primes were calculated for each participant. For the infant version of this analysis, the fitted discriminant analysis classifier was trained on the entire adult dataset and tested on the infant data. D-primes were used to calculate classification accuracy.

3.5 References

- Almeida, J., Fintzi, A. R., & Mahon, B. Z. (2013). Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(9), 2334–2344.
- Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, *20*(10), 1404–1412.
- Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2018). Toddler-Inspired Visual Object Learning. *32nd Conference on Neural Information Processing Systems*. Retrieved from <http://vision.soic.indiana.edu/papers/diversity2018nips.pdf>
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Schmidt, A. M., ... Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454.
- Barrett, T. M., Davis, E. F., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology*, *43*(2), 352–368.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.
- Bushnell, E. W., & Boudreau, J. P. (1993). Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development. *Child Development*, *64*(4), 1005–1021.
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919.
- Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Ralph, M. A. L. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology*, *21*(5), 513–527.
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ... Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*, 13995.
- Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based

- basic-level categories in young infants. *Child Development*, 65(3), 903–917.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, 12(10), 388–396.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598–601.
- Fang, Y., Wang, X., Zhong, S., Song, L., Han, Z., Gong, G., & Bi, Y. (2018). Semantic representation in the white matter pathway. *PLoS Biology*, 16(4), e2003993.
- Fox, M. D. (2018). Mapping Symptoms to Brain Networks with the Human Connectome. *The New England Journal of Medicine*, 379(23), 2237–2245.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171–178.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2017). Neuroanatomical Supplementary Results For A Multi-modal Parcellation of Human Cerebral Cortex. *Nature*, 536, 1–97.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20–25.
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, 15(8), 536–548.
- Guo, Y., Hastie, T., & Tibshirani, R. (2007). Regularized linear discriminant analysis and its application in microarrays. *Biostatistics*, 8(1), 86–100.
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052), 801–806.
- Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458.
- Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The Faces in Infant-Perspective Scenes Change over the First Year of Life. *PloS One*, 10(5), e0123780.
- Kahrs, B. A., & Lockman, J. J. (2014). Building Tool Use From Object Manipulation: A Perception-Action Perspective. *Ecological Psychology: A Publication of the*

- International Society for Ecological Psychology*, 26(1-2), 88–97.
- Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*, 9(1), 133–142.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1476), 2109–2128.
- Lewis, T. L., & Maurer, D. (1992). The development of the temporal and nasal visual fields during infancy. *Vision Research*, 32(5), 903–911.
- Lockman, J. J. (2000). A Perception–Action Perspective on Tool Use Development. *Child Development*, 71(1), 137–144.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., ... Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), 8135–8139.
- Meyer, K., Kaplan, J., Essex, R., Webber, C., Damasio, H., & Damasio, A. (2010). Predicting visual stimuli on the basis of activity in auditory cortices. *Nature Neuroscience*, 13, 667–668.
- Osher, D. E., Saxe, R. R., Koldewyn, K., Gabrieli, J. D. E., Kanwisher, N., & Saygin, Z. M. (2016). Structural Connectivity Fingerprints Predict Cortical Selectivity for Multiple Visual Categories across Cortex. *Cerebral Cortex*, 26(4), 1668–1683.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews. Neuroscience*, 8(12), 976–987.
- Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, 22(4), 463–475.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology: CB*, 27(12), 1825–1828.e3.
- Saygin, Z., & Kanwisher, N. (2014). Structural and functional connectivity fingerprints

- for face, body, scene, and object perception. *Journal of Vision*, *14*(10), 603–603.
- Scheel, A. M., Ritchie, S. J., Brown, N. J. L., & Jacques, S. L. (2017). Through a womb, darkly: Methodological problems in a recent study of fetal visual perception.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology. General*, *133*(2), 166–188.
- Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252.
- Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, *21*(5), 680–685.
- Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, *72*(2), 125–166.

Chapter 4

4 The animacy distinction in the ventral temporal cortex: probing development using vocabulary

Due to the maturity of cortical connectivity in the previous chapter, we decided to further examine cortical development and see if infants were acquiring categories at different rates, in line with general organizing principles in the ventral temporal cortex. In order to measure when infants acquired categories, we used a measure of vocabulary development, termed ‘age of acquisition’, which measures the age at which infants acquire words. In accordance with the animacy distinction in the VTC, we find that infants are acquiring animate words earlier than inanimate words.

4.1 Introduction

In the last chapter I found that in the first nine months of postnatal life, ventral regions processing different visual categories have distinct trajectories of maturation. A prediction from this work, which could not be tested given the data available, is that distinct trajectories of maturation would be reflected in infants’ cognitive capabilities for different categories. In this chapter, we begin to investigate this, by testing if a well-established principle of organization of the ventral temporal cortex (VTC) predicts the sequence in which infants acquire categories. Specifically, in the VTC, there is an animate-inanimate division along the mid-fusiform sulcus (MFS). In fMRI experiments animate stimuli activate one side and inanimate stimuli activate the other (Grill-Spector & Weiner, 2014; Konkle & Caramazza, 2013). Similarly, multivariate pattern classification on VTC activation patterns has found that animate stimuli, primarily faces, body parts and animals, are clustered together, while inanimate stimuli, like fruit, plants, tools, places, and other manmade objects, formed another cluster (Kriegeskorte et al., 2008). The MFS also forms a division in cytoarchitectonics and connectivity profiles, where one side has a different cytoarchitectonic and connectivity profile than the other, that may be necessary for the distinct processing that occurs for each category (Grill-Spector & Weiner, 2014; Osher et al., 2016; Weiner et al., 2014).

Although other distinctions exist within the VTC, we chose to probe the animacy distinction because it is one of the largest scale organizing principles (Grill-Spector & Weiner, 2014). Furthermore, it is fundamental to the inferences we make about the environment. For example, evolutionarily, it is useful to know when something is approaching you and may require increased attention, which contrasts with the processing that would occur in a familiar place; as such, we remember animate items better than inanimate items (Nairne, VanArsdall, & Cogdill, 2017). Infants acquire this distinction quite early in life and it is one of the last distinctions preserved in semantic dementia and Alzheimer's (Hodges, Graham, & Patterson, 1995; Rakison & Poulin-Dubois, 2001). In order to assess the animacy distinction, we probed whether infants produced animate words before inanimate words. In the language literature, this measurement is termed 'age of acquisition' (AoA) and is dissociable from word frequency (Marc Brysbaert, 2017).

Representing the close correspondence between visual stimuli and verbal labels, picture naming tasks are often used to probe AoA. When picture naming tasks are given in adults, items with earlier AoA's have faster reaction times (Bonin, Chalard, Méot, & Fayol, 2002; M. Brysbaert, Van Wijnendaele, & De Deyne, 2000; Ghyselinck, Lewis, & Brysbaert, 2004; Morrison & Ellis, 1995). This has led some researchers to propose a neural network model, where networks that represent these concepts are more plastic earlier on in life--earlier in training--and thus labels that are acquired together may be organized together (Menenti & Burani, 2007).

AoA can be measured in several ways. Participants can be asked retrospectively when they learned a word or their vocabulary can be tested at many time-points through development. Given that word learning continues for more than a decade and there are so many words to test for, it is laborious to measure at each age, and so most studies use estimates of AoA based on retrospective report. However, in our work, we are particularly interested in the earliest stages of development in infancy. These fall within a period of "infantile amnesia" before 3-4 years old, when people universally remember nothing, making it potentially impossible for them to accurately report the age of early word acquisition (Loftus, 1993).

Fortunately, given the clinical importance of detecting delays in language, another type of test is well established in infants, based on caregiver report. A common instrument is the MacArthur-Bates Communicative Development Inventory (MB-CDI), which has been found to have reliability and validity when assessing AoA (Frank, Braginsky, Yurovsky, & Marchman, 2017; Heilmann, Weismer, Evans, & Hollar, 2005; Thordardottir & Weismer, 1996). The MB-CDI groups words into categories. In order to determine what categories should be considered animate and inanimate, we defined animacy using results from previous behavioural experiments. Infants may conceptualize animacy through movement and they may do this differently than adults (Rakison & Poulin-Dubois, 2001). In adults, many researchers define animacy as goal-directed movement, while inanimate objects may move without having a specific goal (i.e. a train moves, but the conductor initiates the movement). However, infants have been found to show a different conceptualization of animacy, failing to distinguish between goal directed and non-goal directed movement. For example, vehicles would fall into an infant's animate category because from their perspective they move in the same way a cat or dog would (Rakison & Poulin-Dubois, 2001). Although this distinction is not often made by adult researchers, there is imaging data to show the organization of semantic categories may reflect it. For example, when exploring the organization of semantic categories Huth et al. (2012) presented adult participants with movies and conducted a principal components analysis. The strongest predictor of the first principle component was movement, with faces, animals and vehicles against other static stimuli (Huth, Nishimoto, Vu, & Gallant, 2012). Indeed, many other studies into the organization of the VTC have failed to include vehicles as a category (Kriegeskorte et al., 2008). Therefore, the definition of animate stimuli was extended to all objects that move. Inanimate stimuli were defined as non-moving objects and included categories similar to those forming clusters in previous research. Based on the division between animate and inanimate stimuli in the VTC, we expect that there might be a differential rate of learning between animate and inanimate categories. Based on the importance of animate stimuli, we might expect animate vocabulary acquisition to lead inanimate vocabulary acquisition.

Additionally, after exploring whether infants acquire animate words before inanimate words, to study the effect of using a parental report measure rather than retrospective

self-report, we compared the MB-CDI parental reports to a database that recorded AoA retrospectively. We expected that the parental reports will have a significantly different AoA because of infantile amnesia.

4.2 Methods

To investigate if infants acquired concepts in accordance with the organization of the ventral temporal cortex (VTC), a corpus analysis was conducted with the data from the Wordbank database (Frank et al., 2017). The data used in this experiment was downloaded from the Wordbank database on November 25th, 2016. The version downloaded stored the data of 5,450 infants who completed the English portion of the MB-CDI. They were between 16 months to 30 months old and were typically developing, exhibiting no form of language disorder, as reported by the researchers who contributed data.

The MB-CDI asks parents to rate whether their child can produce a word. As the MB-CDI is a production measure of language acquisition, it is less correlated with word frequency than other means of assessing AoA (Brysbaert, 2017). As a non-retrospective, reliable, valid tool (Dale, 1991; Heilmann et al., 2005; Thal, O'Hanlon, Clemmons, & Fralin, 1999; Thordardottir & Weismer, 1996) the MB-CDI is a strong way to investigate infant vocabulary development.

Our goal was to investigate whether the animate/inanimate distinction, so prominent in VTC, was predictive of order of acquisition. The words analysed and the categories they were divided into were predetermined by the MB-CDI. However, whether the categories were animate or inanimate was decided based on previous developmental research, where categories that moved were scored as animate and categories that did not moved were scored as inanimate. The categories that were scored as animate were: 'Animals', 'Body parts', 'Vehicles', and 'People' and the categories that were scored as inanimate were: 'Household', 'Places', 'Food and Drink', 'Locations', 'Toys and Furniture', 'Clothing', and 'Outside'. Categories that fell into neither distinction were scored as 'Other', and they were not included in the analysis; the items in this category consisted of words, non-words (e.g. baa-baa) and gestures that were judged to not be relevant. The proportion of

infants who could produce each word was used to determine whether infants acquired animate or inanimate words first, underscoring the distinction seen in the VTC. As the data violated the normality assumption, which is to be expected in language acquisition, we report the median values of each category at the 15 ages contained in the MB-CDI. We then use a non-parametric test, a Wilcoxon sum rank t-test, to determine if infants are producing significantly more animate words than inanimate words, based on the proportion of words produced for each category. We test each age point separately because we wanted to determine the developmental trajectory of the vocabulary development surrounding the animacy distinction.

To investigate the effect of using the MB-CDI rather than a more widely used retrospective measure of AoA, we also examined AoA ratings that were acquired from a database created by Kuperman et al. (2012). Kuperman et al. (2012) calculated mean AoA retrospective ratings using Amazon's Mechanical Turk, an online crowdsourcing platform that allows researchers to gain a diverse sample in a short amount of time for a low cost. This data had a high correlation with data that was acquired by asking participants for the same information in the lab (.93 or .85 depending on the database they compared their data to) (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Participants were asked to report at which age they had learned particular words. This is distinct from the data utilized from the Wordbank database, as in the MB-CDI, production, a more stringent criterion for learning, was the form of reference. As such, if there was a significant difference between the two databases it could be expected that the Kuperman et al. (2012) ratings would trend towards lower ages than what was found in the Wordbank database. In their data collection Kuperman et al. (2012, p. 10) stressed that participants were to report, "the age at which you would have understood that word if somebody had used it in front of you, EVEN IF YOU DID NOT use, read or write it at the time". However, because of infantile amnesia, participants could be reporting significantly higher ages than found for the parental report measure. Kuperman et al. (2012) collected data from 1,960 participants over the course of their study, which gave them the ability to collect data for over 30,000 words.

In order to make comparisons between the data in the MB-CDI and Kuperman's data, which provided a mean AoA, we calculated the mean age at which infants acquired words in the MB-CDI data, by calculating the age where 50% of the infants were producing a word. This threshold was chosen as it is the steepest part of the curve, if the transition is like a sigmoid. We then compared the two AoA measures; there were 389 words scored as being animate or inanimate in the CDI data and corresponding Kuperman AoA's were found for 295 words. Words that did not have clear correspondents (e.g. words that had two noun entries (i.e. chicken (food) and chicken (animal))) were excluded from the analysis. We then tested whether the Kuperman (2012) and MB-CDI AoA for the 295 words were correlated, and if they were significantly different. To probe the animacy distinction, we tested within each dataset whether infant mean AoA was greater for animate or inanimate words, for the subset of 295 words. Again, the data violated the normality assumption and Wilcoxon sum rank t-tests were used.

4.3 Results

MB-CDI Measure of AoA from Wordbank

In order to describe the data, the median proportion of words produced was calculated for the categories that were classified as animate and inanimate. Medians for each category are presented in Table 1. A plot containing the median proportion of words produced for each category and the interquartile range at each age is presented in Figure 10.

Wilcoxon rank sum t-tests were used to assess whether infants were able to produce more animate words than inanimate words. Up until 29 months, infants were significantly able to produce more animate words than inanimate words. Test statistics and significance levels can be found in Table 2.

Table 1. Proportion of words known

Median proportion of words known for animate and inanimate categories

Age (months)	Animate Median	Inanimate Median
16	0.097	0.040
17	0.124	0.047
18	0.212	0.116
19	0.265	0.145
20	0.305	0.178
21	0.345	0.225
22	0.509	0.370
23	0.571	0.440
24	0.628	0.504
25	0.664	0.569
26	0.735	0.641
27	0.726	0.649
28	0.841	0.799
29	0.779	0.759
30	0.858	0.833

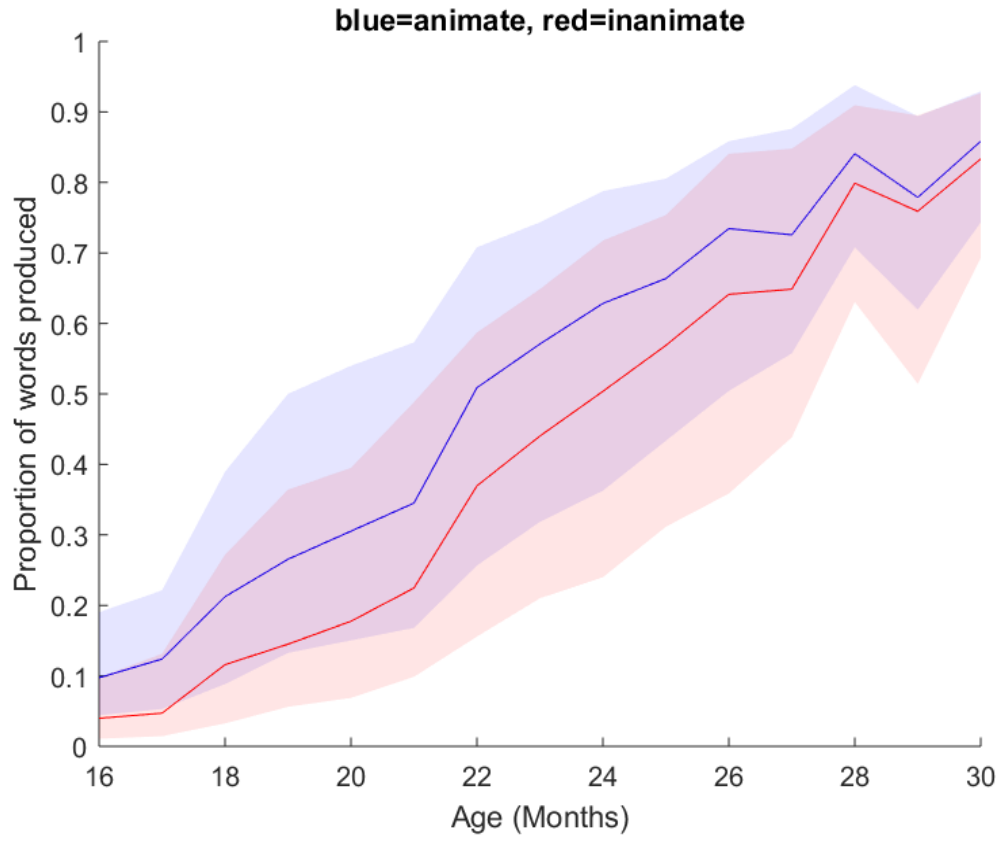


Figure 10. Proportion of words produced for animate (blue) and inanimate (red) categories. Error bars represent the interquartile range.

Table 2. Wilcoxon sum rank tests

Wilcoxon sum rank tests: contrasting the animate and inanimate portion of words learned

Age (months)	Zscore	p value	N
16	15.19	p<0.0001	956
17	7.77	p<0.0001	273
18	7.31	p<0.0001	442
19	6.10	p<0.0001	328
20	5.23	p<0.0001	274
21	3.69	p<0.001	203
22	3.44	p<0.001	198
23	4.2	p<0.0001	258
24	5.49	p<0.0001	555
25	3.43	p<0.001	309
26	2.31	p<0.05	170
27	2.15	p<0.05	195
28	4.87	p<0.0001	796
29	1.51	p>0.05	198
30	1.82	p>0.05	295

Retrospective AoA Ratings

As expected, mean retrospective AoA ratings were substantially different, with Kuperman's ratings being higher (i.e. participants reported learning words later in life) than the mean ratings calculated from the MB-CDIs stored in the Wordbank database (see scatter plot in Figure 11; Wilcoxon rank sum test z-score=-21.01, p<0.00001). Furthermore, the retrospective ratings were not just linearly shifted and scaled, but reflected substantially difference variance, with the correlation between them only of medium strength, despite the high sample sizes in both studies (r=.52, p<0.00001). This shows that retrospective report gives a different sequence.

We then individually examined the effect of animacy, using the mean AoA calculated for the retrospective AoA measure and the MB-CDI. As for the MB-CDI measure, there was a reliable difference between animate and inanimate word mean AoA (Wilcoxon rank sum test zscore=-2.12, p=0.0342). However, the retrospective measure also has a reliable difference between animate and inanimate words (Wilcoxon rank sum test zscore=-2.03,

$p < 0.0395$). A cumulative histogram of the mean proportion of words learned or produced at each age, calculated individually for each database, can be found in Figure 12.

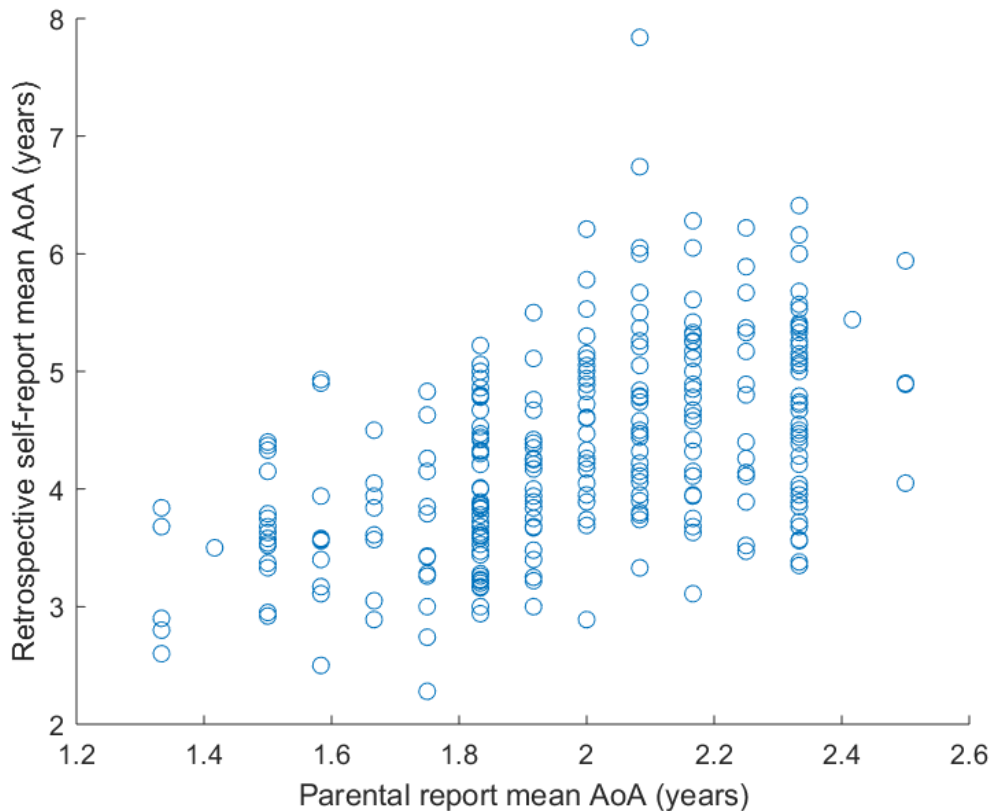


Figure 11. Scatter plot comparing the mean AoA from the parental report measure to the mean AoA obtained from the retrospective self-report measure.

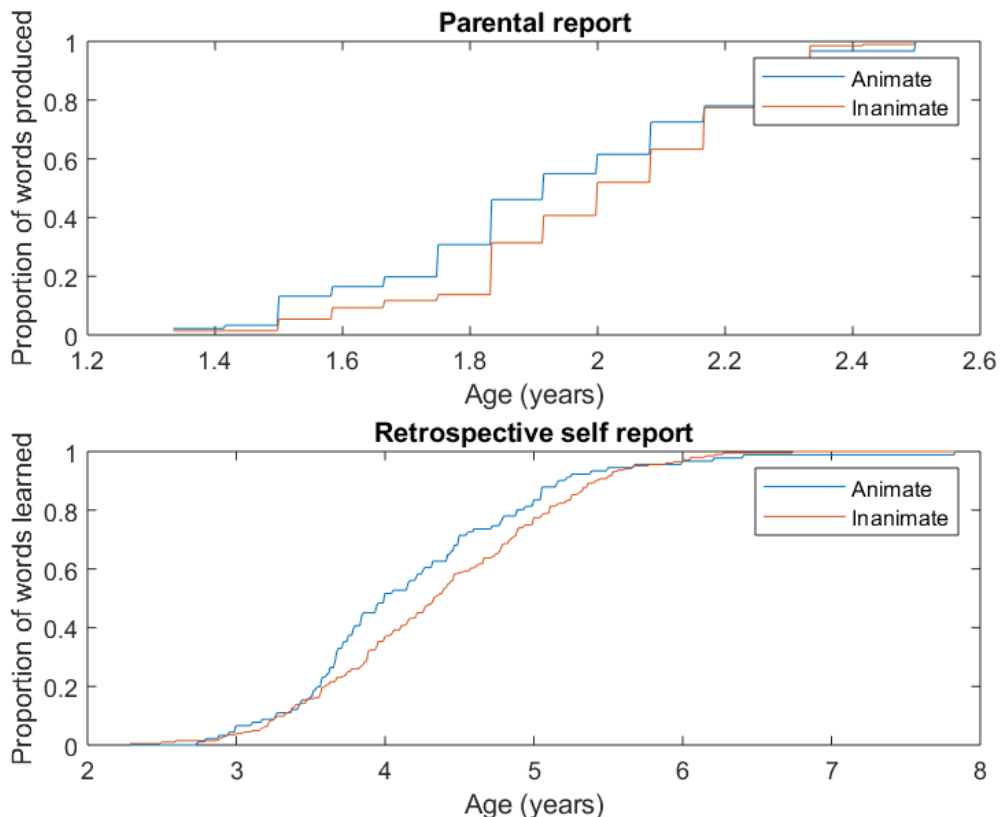


Figure 12. Cumulative histogram for the proportion of words produced (parental report) or learned (retrospective self-report). Animate words are learned earlier in both databases.

4.4 Discussion

In order to determine whether the organization of the VTC was related to the order in which infants acquire categories, we used the Wordbank database which contains MB-CDI data from 5,450 infants to determine if infants learned words from animate categories earlier than inanimate words. As would be expected by the general organizing principle in the VTC, there was a difference in acquisition between animate and inanimate categories. In accordance with the value of these categories, infants acquired animate words before inanimate words.

While this data gives the interesting opportunity to compare the timeline of how categories are acquired with the organization of the VTC, it is still unknown how

vocabulary development influences the maturity of the VTC or vice versa. It is possible that myelination occurs in a prescribed way (possibly laterally to medially) and that the maturity of the VTC influences category development. Animate concepts might be acquired sooner because of the maturity of the lateral portion of the VTC. In this case, it could be said that the cytoarchitectonics and connectivity of the lateral portion of the VTC are genetically coded to represent these stimuli, and the maturity of cortex is needed to acquire concepts. However, it also could be that the visual properties of animate stimuli drive concept acquisition regardless of maturity. The human visual system could be setup to pay attention to things that move, and this could then subsequently drive concept acquisition. In order to definitively identify which comes first, a concurrent neuroimaging and vocabulary development study, measuring infants language production would have to be conducted. This represents a promising avenue for future research.

Although the difference between animate and inanimate word production was no longer significant at 29 months in the MB-CDI data, and the interquartile range had significant overlap, it is possible that this difference continues to be present in older infants and children. The MB-CDI is a fixed questionnaire that might not be capturing all of infant or toddler word development. While there could be real differences in animate and inanimate word knowledge, the infants might be acquiring different animate words at older ages that are not listed on the questionnaire. Future research should seek to use another measurement of language production, such as a tape recorder attached to infants at various ages, to get a greater sense of the words infants are producing. In this case, individual raters would have to code each word produced by the child for a fixed interval for its belonging to an animate or inanimate category.

The data in this study demonstrates that animacy followed a traditional infant definition of animacy, with vehicles being included although their movement is not goal directed in nature (Rakison & Poulin-Dubois, 2001). This aligns with the imaging data from (Huth et al., 2012) and the clustering of other categories within the VTC (Kriegeskorte et al., 2008). Future research should take this into account when measuring the infant animacy distinction, especially if neuroimaging is done in this population.

However, we by no means suggest that the animacy distinction accounts for all of the variance in infant vocabulary development. Therefore, it would be interesting to extend this analysis to other aspects of VTC organization. For example, real object size is also a general organizing principle of the VTC (Konkle & Oliva, 2012). Individual words found in the MB-CDI could be scored to develop two categories: a small object and a large object category. The analysis could then be repeated to determine if infants show this distinction in their vocabulary. Based on the data in this Chapter, we hypothesize that there would be a differential rate of learning between large and small objects. Following the lateral to medial maturation observed here, we hypothesize that small objects, usually clustered on the lateral side of the MFS, would have an earlier AoA than the large objects, usually clustered on the more medial side of the MFS. Additionally, another analysis could look at some of the finer grain distinctions within animacy. For example, in the neuroimaging literature, inanimate items have a further distinction where small inanimate items are cluster together and large inanimate items form another cluster (Konkle & Caramazza, 2013). It is possible that small, inanimate items and large, inanimate items have a differential rates of learning as well.

Although the motivation for using data from the MB-CDI was strong, because we were interested in category development before our first memories are formed, to ensure we were measuring something unique with the Wordbank database, we investigated whether mean AoA's from the MB-CDI and the Kuperman et al. (2012) database were related, finding a medium sized correlation. To determine if retrospectively collecting AoA would influence the comparison between the proportion of animate and inanimate words acquired, we tested the proportion of animate and inanimate words produced in both the Wordbank and Kuperman data individually. Both comparisons were significant. However, the databases were significantly different from each other. This, taken with the medium side correlation, suggests the MB-CDI data could be measuring something unique from the Kuperman data.

This data may be useful for speech language pathologists or other clinicians trying to probe concept development. If infants are not acquiring concepts in accordance with developmental norms, this could have direct implications for both their vocabulary and

the organization of semantics within the ventral visual stream. Clinicians should be aware of this vocabulary difference and that it presents as early as 16 months into postnatal life.

In conclusion, the results provide a promising way to probe infant concept development, by relating the acquisition of verbal labels to VTC organization. Future work should be done to determine the causality of the relationship between the acquisition of verbal labels and the organization of the VTC. This data has clinical implications for both vocabulary and concept development, as neurons that fire together wire together, and abnormal acquisition may shape the organization of the VTC.

4.5 References

- Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(Pt 1), 89–114.
- Brysbaert, M. (2017). Age of acquisition ratings score better on criterion validity than frequency trajectory or ratings “corrected” for frequency. *The Quarterly Journal of Experimental Psychology*, *70*(7), 1129–1139.
- Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*(2), 215–226.
- Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, *34*(3), 565–571.
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: an open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.
- Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: a review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*(1), 43–67.
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, *15*(8), 536–548.
- Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (2005). Utility of the MacArthur—Bates Communicative Development Inventory in Identifying Language Abilities of Late-Talking and Typically Developing Toddlers. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*. Retrieved from <https://ajslp.pubs.asha.org/doi/full/10.1044/1058-0360%282005/006%29>
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: implications for the organisation of semantic memory. *Memory*, *3*(3-4), 463–495.
- Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224.

- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(25), 10235–10242.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114–1124.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., ... Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126–1141.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.
- Loftus, E. F. (1993). Desperately seeking memories of the first few years of childhood: the reality of early memories. *Journal of Experimental Psychology. General*, *122*(2), 274–277.
- Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing? *Quarterly Journal of Experimental Psychology*, *60*(5), 652–660.
- Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *21*(1), 116–133.
- Nairne, J. S., VanArsdall, J. E., & Cogdill, M. (2017). Remembering the Living: Episodic Memory Is Tuned to Animacy. *Current Directions in Psychological Science*, *26*(1), 22–27.
- Osher, D. E., Saxe, R. R., Koldewyn, K., Gabrieli, J. D. E., Kanwisher, N., & Saygin, Z. M. (2016). Structural Connectivity Fingerprints Predict Cortical Selectivity for Multiple Visual Categories across Cortex. *Cerebral Cortex*, *26*(4), 1668–1683.
- Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate-inanimate distinction. *Psychological Bulletin*, *127*(2), 209–228.
- Thal, D. J., O’Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research: JSLHR*, *42*(2), 482–496.
- Thordardottir, E. T., & Weismer, S. E. (1996). Language assessment via parent report: development of a screening instrument for Icelandic children. *First Language*,

16(48), 265–285.

Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., ... Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, *84*, 453–465.

Chapter 5

5 Discussion.

5.1 General Discussion and Conclusions

It has been proposed that infants primarily use a subcortical retinocollicular pathway for detecting a basic level category, faces (Johnson, Senju, & Tomalski, 2015; Francesca Simion, Valenza, Umiltà, & Barba, 1998). This pathway has also been proposed to facilitate rapid detection in adults, and so I began by investigating this with a face detection task. I exploited the greater connectivity of the nasal hemiretina to the retinocollicular pathway, compared to the temporal hemiretina (Bertini, Leo, & Làdavas, 2008; Sylvester, Josephs, Driver, & Rees, 2007). Across four experiments, which manipulated the coarseness and spatial frequency of the stimuli, I failed to find an advantage of presenting to the nasal hemiretina, thus demonstrating a lack of involvement of subcortical structures in rapid face detection in adults. As I was conducting these experiments, evidence emerged of cortical processing of faces, places and objects in the infant VTC (Deen et al., 2017). I therefore focused upon characterizing cortical systems in early infancy, and in particular the development of the unique signatures of connectivity for category-selective regions. I found that the face and place network had mature connectivity early in infancy, while the tool network took nine months of postnatal life to reach the maturity level of adults. These unique signatures of connectivity allow different brain regions come together to make up the cross-modal, motoric and affective associations we associate with rich representations of categories (Binder, Desai, Graves, & Conant, 2009; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Patterson, Nestor, & Rogers, 2007). Given the evidence of maturation within the VTC, I hypothesized that the principles of organization of the VTC might be evident in the order in which infants acquire categories. To determine when infants acquired categories, I used a measure of language production, the Mac-Arthur Bates Developmental Inventory (CDI), a parental questionnaire about their child's vocabulary development. I used the Wordbank database, which compiled thousands of CDI's across a variety of ages (Frank, Braginsky, Yurovsky, & Marchman, 2017). As it is one of the largest scale category distinctions in VTC, I chose to probe the animacy distinction,

delimited by the mid-fusiform sulcus (Weiner et al., 2014). It is also one of the last distinctions to be eliminated in semantic dementia (Hodges, Graham, & Patterson, 1995). Additionally, moving stimuli capture infants' attention (McKenzie & Day, 1976; Volkman & Dobson, 1976), which lead me to hypothesize that infants would acquire animate words before inanimate words. In line with the hypothesis, infants produced significantly more animate words than inanimate words up until 29 months. Additionally, to validate my choice of language measurement, I compared the data from the MB-CDI to the data from a database that collected AoA measurements through retrospective self-report (Frank et al., 2017; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). The mean AoA's between the two databases were reliably different from each other, and the correlation between them was only of medium strength (.52), demonstrating I was measuring something unique with the MB-CDI.

Overall, the work contained in this dissertation demonstrates the surprising maturity of the cortical networks for vision in infants. Even in adults, subcortical regions are making less of a contribution than originally thought. Infants had mature signatures of connectivity, sometimes at only a month old, and the organization of the VTC was related to the order in which infants acquire categories.

However, considering Chapters 3 and 4 together, the question arises of why, if infants are forming categories with rich associations even early in the first year, why does language not develop until the second year? One possibility here concerns the level of description. It could be that the representations were seen to develop in Chapter 3 are either at a more general or more specific than those at the basic level probed through words in Chapter 4. For example, infants may initially make very crude distinctions like animate/inanimate, or be separating agents from non-agents (Kinzler & Spelke, 2007). Or, at a much more specific level, they may know the faces of their immediate family at a month old and be building associations with these, but may yet to generalise relevant associations to all faces, which is perhaps necessary for a basic-level category representation to form. Likewise, an infant may only know a few places, but constant passive perception of places may be enough to drive the connectivity of the place network. Finally, tool-network connectivity might take time to set up, until children can manipulate objects, but

not every tool or object might need to be individually represented to drive network connectivity. Indeed, plasticity for all of these categories continues into adulthood, with the ability to learn and dissociate between new faces, places, and objects or tools.

Additionally, even when disregarding the possibility of a Type 2 error, the lack of evidence for subcortical involvement, via the retinocollicular pathway, in rapid face detection does not mean that subcortical structures are not involved in other face or category level processing. The amygdala has been shown to be activated when viewing both emotional and neutral faces (Mende-Siedlecki, Verosky, Turk-Browne, & Todorov, 2013). Its role in face processing might be the emotional appraisal of the stimulus, which could contribute to memory formation as in the case of flashbulb memories (Adolphs, Cahill, Schul, & Babinsky, 1997; Akirav & Richter-Levin, 2006; Kensinger, Addis, & Atapattu, 2011). Additional research has shown a monocular advantage to face processing that was not evident for other categories, where they hypothesized that monocularly presenting stimuli resulted in more subcortical involvement (Gabay, Burlingham, & Behrmann, 2014). In this study, the researchers tried to present preferentially to the nasal and temporal hemiretina, but also did not find an effect. The lack of neuroimaging in this study makes it hard to assess what subcortical structures are involved, but the researchers hypothesized that the lateral geniculate nucleus played a role. Completing a similar experiment while participants also viewed stimuli monocularly, using MRI compatible goggles, in the MRI scanner would give insight into what, if any, subcortical structures are involved.

Although my data represents a significant first step in identifying the maturity and contribution of different brain networks and their regions to categorization, there are limitations to this research. For example, in Chapter 2, which contained the four experiments that probed the contribution of the retinocollicular pathway, while the stimuli were presented too quickly for participants to make a saccade, in an attempt to ensure the presentation was made to the correct hemiretina, there was no eye tracking in this study. Although participants were instructed to maintain fixation, eye tracking would allow us to be confident that they complied with the instructions. Future research could

replicate my results with eye tracking to have full certainty that participants were maintaining fixation.

In Chapter 3, we defined the category-specific regions using the localizers from the Human Connectome Project, which contained 210 participants. Thus, our choice of category-specific regions well-powered and likely to be reliable at the group level. However, as mentioned, it would have been useful to have functional localizers and diffusion imaging data in the same adult and infant participants. In adults, this might have allowed for more precision in training the classifier, as I would have been able to identify the face, place, and tool selective regions in individual participants, where the exact border of these varies among individuals. It is possible that this could have improved training and subsequent testing in the infants. Future research could investigate and test this.

In the infant participants, having both localizers and diffusion data could have allowed us to see if the regions that were showing the distinct signature of connectivity of the face, place and tool network were specializing functionally as well. This would have allowed me to determine the order in which functional selectivity is related to connectivity, where functional selectivity could precede connectivity or vice versa. However, the difficulty is that the visual localizers would have required the infants to be awake, but the diffusion acquisition required infants to be asleep so that they were sufficiently still. Thus, the exact relationship between functional selectivity and connectivity remains a topic for challenging future research.

The developmental relationship between acquiring concepts and the organization of the VTC remains somewhat ambiguous in Chapter 4. Although I found that infants acquired animate words before inanimate words, the direction of causality between this and the organization of VTC is unclear. Does the organization of the VTC influence concept acquisition, or does concept acquisition drive the organization of the VTC? Future longitudinal research probing concept development through vocabulary concurrently with imaging, and a method like latent change score analysis might disentangle these (Kievit et al., 2018).

Furthermore, it would be interesting to use the methodology from Chapter 4 to probe for other distinctions in early infant vocabulary development. The animacy distinction clearly does not account for all of the variance in infant vocabulary development, and it is very likely that additional distinctions are evident. For example, infants could acquire small inanimate categories before large inanimate categories. Infants might have more experience with smaller objects, as they could interact with them and hold them up to their faces, making them more likely to acquire those concepts.

Additionally, it would be interesting to probe whether individual differences in infants' frequency of experience with different visual categories drives their vocabulary development and the subsequent organization of the VTC. For example, during eating, it has been shown that infants acquire words that they have experience with at meal time earlier in development, contrasting with words they do not have experience with (Clerkin, Hart, Rehg, Yu, & Smith, 2017). It might be that mugs and staplers are clustered together in a particular infant's VTC if they have experience with both of those items at the same time, whereas another child could acquire candles and books if they were being read to at the dinner table. Substantiating this hypothesis, children (age 5-8) who had significant experience with Pokemon developed a region in the VTC that uniquely responded to Pokemon (Gomez, Barnett, & Grill-Spector, 2019). Further research is needed to determine the relationship between early experience and the organization of the VTC.

Additional research could also be done with adult participants to illuminate the relationship between word AoA and the organization of the VTC. Specifically, it would be interesting to test if the order of concept acquisition in infancy – as reflected in AoA, and measured by the Wordbank database – is predictive on a finer scale of organization within the mature VTC. In other words, are items “laid down in order”. Database data could be used to determine responsiveness in the adult VTC; for example, responsiveness to individual images could be measured using the BOLD 5000 database (<https://bold5000.github.io/>). A univariate, continuous analysis could be conducted to determine whether the images presented are represented differently depending on the average age they were acquired (as dependent on the Wordbank Database). Based on the results in Chapter 4, and adult data that shows the animate/inanimate distinction in the

VTC, I would hypothesise that infants' age of acquisition will be at least somewhat predictive of VTC organization (Grill-Spector & Weiner, 2014). Using a database like the Bold 5000 database would also allow me to probe whether additional factors are predictive of VTC organization through a multivoxel pattern analysis. These factors include, but are not limited to, real objects size, eccentricity, animacy, living vs non-living and combinations of the these.

Furthermore, I could use the same methods developed in Chapter 3 to examine whether regions that are selective for the early and late acquired words have adult-like structural connectivity. By splitting the words into different groups, I could then create ROI's within the VTC for early and late learned words. I would hypothesise that regions that represent words acquired earlier in infancy will have connectivity that is more similar to the adult connectivity when infants are young (i.e. if infants have acquired a set of words represented in a particular region, it is more likely that their brain connectivity will look like adults for that region, whereas the opposite should be true for regions that represent words that are acquired later in infancy), demonstrating that language and a full network representation develop at similar time points.

The methods used in Chapter 3 also have many applications for the study of infant brain development. To my knowledge, our study is the first to use a classifier, trained on adult and tested on infant data, to characterize the connectivity of a brain region. Future research could explore other brain regions, such as those involved in executive function, to determine if the connectivity related to those regions is more mature than previously thought (Rothbart, Sheese, & Posner, 2007). Future research could also explore regions related to theory of mind, which are thought to take significant time to develop (Saxe, Carey, & Kanwisher, 2004).

These methods might also be valuable in determining if participants are at risk for a variety of disorders that involve structural connectivity. Structural connectivity has been used to predict the severity of several health conditions, including mental health conditions (i.e. social anxiety disorder) (Whitfield-Gabrieli et al., 2016). Furthermore, infants with preterm birth and autism have been found to have abnormal preferences for

different visual categories (Adolphs, Sears, & Piven, 2001; Kleinhan et al., 2008; Telford et al., 2016) and it would be interesting to evaluate our signature of connectivity method, to see if it can act as an earlier biomarker for the detection of disorders. Early identification would allow for potentially more effective interventions, when the brain is most plastic. Identification of risk factors during infancy would allow parents to participate in interventions and develop parenting and coping skills that would be useful in providing the best environment possible for infants to grow up in.

Overall, my work demonstrates the role and maturity of cortical structures and brain connectivity in categorization. It developed many methods, some using computational techniques from machine learning, that will be useful for studying the manner in which experience shapes the organization of the VTC, and methods that will be useful for examining brain connectivity, both throughout healthy development and in many disorders that involve structural connectivity. One advantage to the method developed in Chapter 3 is that it allows researchers to meaningfully compare between groups. For example, term infants could be compared to preterm infants to examine typical development, patients could be compared with controls, and infant connectivity could continue to be contrasted with adult connectivity. Our method was specific enough to examine how individual infants compared with individual adults, making it a promising avenue to identify developmental abnormalities on an individual subject level.

My work also underscores the utility of publicly available data when studying infant development. Publicly available databases provided power and generalizability to my work, and collecting all of the data individually would have required significant effort. Future research should seek to exploit the advantages of publicly available data whenever possible—there are clearly many unique questions that databases lend themselves to. Finally, my work attempted to link behaviour and brain function during vocabulary acquisition. Knowing how the infant brain functions is interesting, but greater advances in the field will be made by linking brain development with behaviour, including disordered behaviour that continues to develop throughout infancy and into adulthood.

Taken together, this work presents a basis for many promising future studies. Future work should especially pursue the study using the BOLD 5000 database to gain insight into how the VTC develops. However, I think one of the most important contributions are the methods that have been developed. The VTC represents an interesting development ground for methods because its cognitive functions are partially understood, and because its organization is in some ways robust across participants. The methods described in Chapter 3 are an exciting path to investigate whether the signature of connectivity for particular regions involved in disorders is differentiated between groups and on an individual subject level. To increase power and generalizability, publicly available data should be used whenever possible. Finally, to move the field forward, neuroimaging and behavioural tasks should to be completed with the same participants, allowing for the examination of brain and behaviour interactions.

5.2 References

- Adolphs, R., Cahill, L., Schul, R., & Babinsky, R. (1997). Impaired declarative memory for emotional material following bilateral amygdala damage in humans. *Learning & Memory*, *4*(3), 291–300.
- Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience*, *13*(2), 232–240.
- Akirav, I., & Richter-Levin, G. (2006). Factors that determine the non-linear amygdala influence on hippocampus-dependent memory. *Dose-Response: A Publication of International Hormesis Society*, *4*(1), 22–37.
- Bertini, C., Leo, F., & Làdavas, E. (2008). Temporo-nasal asymmetry in multisensory integration mediated by the Superior Colliculus. *Brain Research*, *1242*, 37–44.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.
- Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711).
<https://doi.org/10.1098/rstb.2016.0055>
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ... Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*, 13995.
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: an open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.
- Gabay, S., Burlingham, C., & Behrmann, M. (2014). The nature of face representations in subcortical regions. *Neuropsychologia*, *59*, 35–46.
- Gomez, J., Barnett, M., & Grill-Spector, K. (2019). Extensive childhood experience with Pokémon suggests eccentricity drives organization of visual cortex. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-019-0592-8>
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral

- temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, *15*(8), 536–548.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: implications for the organisation of semantic memory. *Memory*, *3*(3-4), 463–495.
- Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458.
- Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neuroscience and Biobehavioral Reviews*, *50*, 169–179.
- Kensinger, E. A., Addis, D. R., & Atapattu, R. K. (2011). Amygdala activity at encoding corresponds with memory vividness and with memory for select episodic details. *Neuropsychologia*, *49*(4), 663–673.
- Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., ... Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience*, *33*, 99–117.
- Kinzler, K. D., & Spelke, E. S. (2007). Core systems in human cognition. In C. von Hofsten & K. Rosander (Eds.), *Progress in Brain Research* (Vol. 164, pp. 257–264). Elsevier.
- Kleinhans, N. M., Richards, T., Sterling, L., Stegbauer, K. C., Mahurin, R., Johnson, L. C., ... Aylward, E. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain: A Journal of Neurology*, *131*(Pt 4), 1000–1012.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.
- McKenzie, B. E., & Day, R. H. (1976). Infants' attention to stationary and moving objects at different distances. *Australian Journal of Psychology*, *28*(1), 45–51.
- Mende-Siedlecki, P., Verosky, S. C., Turk-Browne, N. B., & Todorov, A. (2013). Robust selectivity for faces in the human amygdala in the absence of expressions. *Journal of*

- Cognitive Neuroscience*, 25(12), 2086–2106.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews. Neuroscience*, 8(12), 976–987.
- Rothbart, M. K., Sheese, B. E., & Posner, M. I. (2007). Executive Attention and Effortful Control: Linking Temperament, Brain Networks, and Genes. *Child Development Perspectives*, 1(1), 2–7.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55, 87–124.
- Simion, F., Valenza, E., Umiltà, C., & Barba, B. D. (1998). Preferential orienting to faces in newborns: A temporal--nasal asymmetry. *Journal of Experimental Psychology. Human Perception and Performance*, 24(5), 1399.
- Sylvester, R., Josephs, O., Driver, J., & Rees, G. (2007). Visual fMRI responses in human superior colliculus show a temporal-nasal asymmetry that is absent in lateral geniculate and visual cortex. *Journal of Neurophysiology*, 97(2), 1495–1502.
- Telford, E. J., Fletcher-Watson, S., Gillespie-Smith, K., Pataky, R., Sparrow, S., Murray, I. C., ... Boardman, J. P. (2016). Preterm birth is associated with atypical social orienting in infancy detected using eye tracking. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 57(7), 861–868.
- Volkman, F. C., & Dobson, M. V. (1976). Infant responses of ocular fixation to moving visual stimuli. *Journal of Experimental Child Psychology*, 22(1), 86–99.
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., ... Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, 84, 453–465.
- Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, 21(5), 680–685.

Appendices

Appendix A: Regions Driving Connectivity

In order to determine which regions carried the most signal for classification, the mean difference in connectivity between category-selective and non-category selective seed voxels was calculated for all target regions. It was standardized across target regions to yield a z-score. This table shows regions with a z-score that is greater than 3.

Face

Name	Region Number	Z score
PH	138	4.1
TE2 posterior	136	4.06
TE1 posterior	133	3.51
Area PGp	143	3.35
Area PGs	151	3.07

Place

Name	Region Number	Z score
Hippocampus	120	4.59
ParaHippocampal Area 3	127	4.5
Entorhinal Cortex	118	3.84
ParaHippocampal Area 2	155	3.82
Perirhinal Ectorhinal Cortex	122	3.45
ParaHippocampal Area 1	126	3.01

Tool

Name	Region Number	Z score
Fourth Visual Area	6	5.08
Third Visual Area	5	4.07
Area TG dorsal	131	3.25
Pirform Cortex	110	3.07
PosteriorOFC Complex	166	3.03

Appendix B: Ethics Approval



**Western
Research**

Western University Health Science Research Ethics Board HSREB Amendment Approval Notice

Research Ethics

Principal Investigator: Dr. Rhodri Cusack
Department & Institution: Social Science/Psychology, Western University

HSREB File Number: 103665
Study Title: Developing assessments of perinatal brain injury using fMRI
Sponsor:

HSREB Amendment Approval Date: November 21, 2014
HSREB Expiry Date: April 30, 2018

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Revised Western University Protocol	Received Nov 14, 2014	
Revised Letter of Information & Consent	controls	2014/11/14
Revised Letter of Information & Consent	NAS participants	2014/11/14
Revised Letter of Information & Consent	Patients	2014/11/14
Instruments	Drug use questionnaire	2014/11/14

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the amendment to the above named study, as of the HSREB Amendment Approval Date noted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review. If an Updated Approval Notice is required prior to the HSREB Expiry Date, the Principal Investigator is responsible for completing and submitting an HSREB Updated Approval Form in a timely fashion.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice Practices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.



Ethics Officer to Contact for Further Information

<input checked="" type="checkbox"/> Erika Basile ebasile@uwo.ca	<input type="checkbox"/> Grace Kelly grace.kelly@uwo.ca	<input type="checkbox"/> Mna Mkhail mmekhail@uwo.ca	<input type="checkbox"/> Vikki Tran vikki.tran@uwo.ca
--------------------------------------------------------------------	------------------------------------------------------------	--------------------------------------------------------	----------------------------------------------------------

This is an official document. Please retain the original in your files.



**Western University Health Science Research Ethics Board
HSREB Annual Continuing Ethics Approval Notice**

Date: February 28, 2015
Principal Investigator: Dr. Rhodri Cusack
Department & Institution: Social Science/Psychology, Western University

HSREB File Number: 101988
Study Title: Neural mechanisms of visual and auditory perception, short-term memory, and imagery - 18625E
Sponsor:


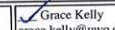


HSREB Renewal Due Date & HSREB Expiry Date:
 Renewal Due -2016/01/31
 Expiry Date -2016/02/02

The Western University Health Science Research Ethics Board (HSREB) has reviewed the Continuing Ethics Review (CER) Form and is re-issuing approval for the above noted study. The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice (ICH E6 R1), the Ontario Freedom of Information and Protection of Privacy Act (FIPPA, 1990), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Ethics Officer to Contact for Further Information

 Erika Basile ebasile@uwo.ca	 Grace Kelly grace.kelly@uwo.ca	 Mna Mekhail mmekhail@uwo.ca	 Vikki Tran vikki.tran@uwo.ca
---------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------

This is an official document. Please retain the original in your files.



**Western
Research**

Research Ethics

**Western University Non-Medical Research Ethics Board
NMREB Annual Continuing Ethics Approval Notice**

Date: December 01, 2016

Principal Investigator: Dr. Rhodri Cusack

Department & Institution: Social Science\Psychology, Western University

NMREB File Number: 105882

Study Title: Nature of visual and auditory representations during perception, memory, and imagery

Sponsor: Natural Sciences and Engineering Research Council

NMREB Renewal Due Date & NMREB Expiry Date:

Renewal Due -2017/11/30


Expiry Date -2017/12/10

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed the Continuing Ethics Review (CER) form and is re-issuing approval for the above noted study.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), Part 4 of the Natural Health Product Regulations, the Ontario Freedom of Information and Protection of Privacy Act (FIPPA, 1990), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.


Ethics Officer, on behalf of Dr. Riley Hinson, NMREB Chair

Ethics Officer: Erika Basile ___ Katelyn Harris ___ Nicole Kaniki Grace Kelly ___ Vikki Tran ___ Karen Gopaul ___

Curriculum Vitae

Education

- 2013/09-
Current PhD *Psychology*, Brain & Mind Institute, Western University, ON,
Canada
Transferred from Master's in 2015.
Supervised by: Rhodri Cusack and Jody Culham.
- 09/09-13/04 BA (Hons) *Psychology*, Huron University College at Western
University, ON, Canada

Memberships

Society for Neuroscience
Canadian Society for Neuroscience
Organization for Human Brain Mapping
International Congress on Infant Studies

A. Peer Reviewed Journal Articles

In Preparation

- A5. Chiara C., **Cabral, L.**, Cusack, R. (In Prep). Executive network connectivity in infancy.
- A4. **Cabral, L.**, Turner, K., Cusack, R. (In Prep). The animacy distinction: Vocabulary development in infancy.

Submitted

- A3. **Cabral, L.**, Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (Submitted). Category-Selective visual regions have distinctive connectivity in early infancy.
- A2. **Cabral, L.**, Stojanoski, B., Cusack, R. (Submitted). Rapid and coarse face detection in cortex.

2017

- A1. Tran, M., **Cabral, L.**, Patel, R., Cusack, R. (2017). Online recruitment and testing of infants with Mechanical Turk, *Journal of Experimental Child Psychology*, 156, 168-178.

C. Conference Abstracts

- C11. Cabral, L., Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (2018). “Visual Categories in infants: local selectivity and distributed connectivity mature together” *Sixth Biennial Conference on Resting State and Brain Connectivity*.
- C10. Caldinelli, C., Cabral, L., Cusack, R. (2018). “Smart on the Inside: Functionally Rich Differences in Connectivity Within the Dorsolateral Prefrontal Cortex Are Present in Early Infancy” *Conference on Cognitive Computational Neuroscience, Accepted Talk*.
- C9. Cabral, L., Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (2018.) “Development of face, place, and object recognition systems in infants: Local selectivity and distributed networks emerge together” *International Congress on Infant Studies*.
- C8. Cabral, L., Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (2018). “Visual Categories in infants: local selectivity and distributed connectivity mature together” *Human Brain Mapping*.
- C7. Cabral, L., Zubiaurre, L., Wild, C., Linke, A., Cusack, R. (2017). “Category-Selective visual regions show distinctive signatures of connectivity in early life” *Society for Neuroscience*.
- C6. Cabral, L., Stojanoski, B., Cusack, R. (2015). “Rapid, accurate detection of faces, but not houses suggests a visually coarse, but context sensitive subcortical face processing mechanism” *Society for Neuroscience*.
- C5. Cabral, L., Stojanoski, B., Cusack, R. (2015). “Rapid and Coarse Face Detection in Cortex” *Inaugural Brain and Mind Conference*.
- C4. Tran, M., Cabral, L., Cusack, R. (2015). “Dora the Neurocognitive Explorer: Identifying engaging movies for infant fMRI” *Inaugural Brain and Mind Conference*.
- C3. Cabral, L., Stojanoski, B., Wild, C., Cusack, R. (2015). “Investigating Cortical and Subcortical Influences on Face Perception” *Lake Ontario Visionary Establishment Conference*.
- C2. Cabral, L., Stojanoski, B., Wild, C., Cusack, R. (2013). “In search of babies: Using Mechanical Turk to examine infant stimulus preferences” *Canada-Israel Symposium on Brain Plasticity, Learning, and Education*.
- C1. Cabral, L., Stojanoski, B., Wild, C., Cusack, R. (2013). “In search of babies: Using Mechanical Turk to examine infant stimulus preferences” *Canadian Association for Neuroscience*.

Scholarships

Natural Sciences and Engineering Research Council Postgraduate Scholarship-
 Doctoral Award (2017-2019), \$21,000/yr.
 Ontario Graduate Scholarship (2015-2016), \$15,000/yr.
 Canadian Institute of Health Research Master's Scholarship (2014-2015), \$17500/yr.
 Ontario Graduate Scholarship [declined] (2014), \$15,000
 Frank P. Leahy Scholarship (2011), \$3,000.
 Huron Scholarship (2009-2012), \$2000/yr.

Prizes and Awards

Nominated for King's Undergraduate Teaching Award (2018-2019)
 Travel Award (2017) Society for Neuroscience
 First Place Business Case Competition (2017) Ivey Mini MBA GMCA, Western University
 Experimental Psychology Award (2012) Huron University College, Psychology.
 Chinese Political Science Book Prize (2012) Huron University College, Political Science.
 Quantitative Analysis Award (2011) Huron University College, Psychology.

Supervision

Kristen Turner, 2017, Psychology, Honors Thesis, Western University, London, ON,
 Supervised with Rhodri Cusack

Teaching

Lecturer, Special Topics in Cognitive Neuroscience, Third Year Undergraduate, 3994G,
 Department of Psychology, King's University College at Western, ON, Canada (2019).
 Teaching Assistant, Introduction to Developmental Psychology, Second Year
 Undergraduate, 2410B, Department of Psychology, Western University, ON, Canada
 (2018).
 Teaching Assistant, Memory, Third Year Undergraduate, 3233G, Department of
 Psychology, Western University, ON, Canada (2017).
 Teaching Assistant, The Emergence of Cognition, Third Year Undergraduate, 3295F,
 Department of Psychology, Western University, ON, Canada (2016).
 Teaching Assistant, The Emergence of Cognition, Third Year Undergraduate, 3295F,
 Department of Psychology, Western University, ON, Canada (2015).
 Teaching Assistant, Current Topics in Neuroscience, Third Year Undergraduate,
 N3000G, Neuroscience Program, Western University, ON, Canada (2015).
 Teaching Assistant, Tests and Measurement, Second Year Undergraduate, 2002A,
 Department of Psychology, Western University, ON, Canada (2014).
 Teaching Assistant, Introduction to Psychology, First Year Undergraduate, 1000,
 Department of Psychology, Western University, ON, Canada (2013-2014).