Title
Representing Geometry: Perception, Concepts, and Knowledge

Author
Galebach, John Ethan

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Representing Geometry: Perception, Concepts, and Knowledge

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Philosophy

by

J. Ethan Galebach

Dissertation Committee:
Professor Jeremy Heis, Chair
Professor Penelope Maddy
Professor Sean Walsh

2018

TABLE OF CONTENTS

Page

Acknowledgments **iv**

Curriculum Vitae **vii**

Abstract of the Dissertation **viii**

Introduction **1**

1 Does Visual Perception Have a Geometry? **5**
1.1 Introduction 6
1.2 VMS Proponents and Arguments 8
1.2.1 What is Visual Perception? 8
1.2.2 Five Arguments for a VMS 11
1.2.3 Behavioral evidence against a VMS 26
1.3 Spatial Representation according to Vision Science 30
1.3.1 Marr's visual perception: demarcation and primary function 30
1.3.2 Current Views on Shape and Scene Perception 36
1.4 Conclusion and Open Questions 43
1.5 Bibliography 44

2 Are Euclidean Beliefs Grounded in Core Cognitive Systems? **57**
2.1 Introduction 57
2.2 The Core Cognitive Theory of Numerical Concept Acquisition 58
2.2.1 Two Core Systems of Number 59
2.2.2 Bootstrapping to Numerical Concepts and Arithmetical Beliefs 62
2.3 The Core Cognitive Theory of Geometric Concept Acquisition 65
2.3.1 Two Core Systems of Geometry 67
2.3.2 Bootstrapping to Geometric Concepts and Beliefs 70
2.4 Problems for the Core Cognitive Theory of Geometry 71
2.4.1 Critique of the Two Core Systems of Geometry 71
2.4.2 Critique of the Proposed Geometric Concepts and Beliefs 74
2.5 Bibliography 76

3 Visual Imagination and Mathematical Entitlement **82**
3.1 Introduction 82
3.2 Imagery-based Mathematical Entitlement 85
3.3 The Geometric Contents of Visual Imagination 88
3.4 Geometric Knowledge through Visual Imagination 91
3.5 Knowing the o-minimal IVT through Visual Imagination 95
3.5.1 The Synthetic IVT 96
3.5.2 The Analytic IVT 97
3.5.3 The O-minimal IVT 100
3.6 O-minimality as a Visual-Epistemological Project 108
3.6.1 Can o-minimal objects be visualized? 108
3.6.2 Visual Virtues of O-minimality 112
3.7 Appendix: Modal Logic of O-minimal Structures 118
3.8 Bibliography 122

ACKNOWLEDGMENTS

This project would not have been possible without the encouragement, guidance, and expertise of dozens of philosophers. I want to thank these people, and especially my advisor Jeremy Heis, for developing my scholarship over many years and for treating my whole self with compassion and respect at every turn. In addition, Pen Maddy and Sean Walsh have dedicated the most time and effort to my academic and personal development. I am deeply grateful for the impact they have had on me. Below, I would like to describe my academic development over the last decade in order to thank the many other charitable souls who have contributed to the successful completion of this project.

In my undergraduate studies at the University of Pittsburgh, Mark Wilson was the first professor to show a sincere interest in my philosophical questions. Without Mark's sustained attention to my senior thesis over countless one-on-one sessions in his office, I doubt I would have developed the confidence and self-worth required to articulate and defend a PhD thesis. I hold a similar appreciation for the advice and generosity of Bob Batterman and Tom Ricketts. In my first two years of graduate school at UC Irvine, I was fortunate enough to take multiple seminars with Jeremy, Pen, and Sean on a variety of topics within the philosophy of mathematics. Although the subject matters of these seminars varied widely (e.g., Cassirer's philosophy of geometry, Heyting arithmetic, Dedekind's conceptual analysis of continuity, Freges notion of analyticity), each seminar was intensely concerned with documenting how philosophical questions arise naturally out of mathematical practice. I was absorbed by their treatment of mathematics as a *science* (with its own anomalies, methodological commitments, and historical development). These seminars led me to books like Kitcher's *The Nature of Mathematical Knowledge*, Lakatos' *Proofs and Refutations*, Mancosu's (ed.) *Philosophy of Mathematical Practice*, Bos' *Redefining Geometrical Exactness*, and Giaquinto's *Visual Thinking in Mathematics*, all of which seemed to me to be asking and answering the most interesting and important questions regarding the nature of geometric knowledge.

In the summer of 2014, I became more confident in the legitimacy and durability of these approaches to the philosophy of mathematics when I participated in the month-long PhilMath Intersem in Paris organized by Mic Detlefsen. While I was there I presented a paper on Giaquinto's theory of mathematical entitlement (which eventually became Chapter 3) and I had dozens of invaluable meal-length conversations with philosophers and historians of mathematics from a handful of countries. In particular, Andrew Arana, Karine Chemla, Mic Detlefsen, Katherine Dunlop, Renaud Chorlay, Paolo Mancosu, David Rabouin, and Vincenzo De Risi all displayed genuine interest in my research and they each offered meaningful advice and encouragement for my academic future. My exposure to this constellation of ideas and well-rounded academics solidified my commitment to studying the history of geometry and the nature of spatial representation. Three months later, I presented a logic-focused variant of my paper on Giaquinto at another interdisciplinary conference called *Spatial Cognition* in Riga, Latvia. This is where I met Nora Newcombe, Barbara Landau, and many other psychologists and linguists of spatial representation. They introduced me to an adjacent world of concepts, theorists, and debates that inform all of my chapters. The Spatial Cognition conference made me aware of the close alignment between philosophical and psychological questions about the nature of spatial representation. Giaquinto's book was merely the tip of an iceberg. When I returned, Pen Maddy introduced me to Stephen Palmers *Vision Science* and David Marr's *Vision*, both of which became my portals into the debates and assumptions driving various traditions within cognitive science.

The most general question of my dissertation had taken shape: what are the prospects for using experimental and theoretical psychology to answer common philosophical questions about geometric representation? In terms of isolating questions within the philosophical discourse on spatial perception, Pen's seminar on vision theory in the Winter of 2016 introduced me to the work of Gary Hatfield, and a reading group with Jeremy and Alysha Kassam in Summer of 2016 introduced me to the work of Tyler Burge. The conversations about the border of perception and whether its content had a metric relational structure arose out of

these two experiences. Chapter 1 was born. Regarding concept acquisition, a handful of LPS colloquia on arithmetical knowledge (as well as the works of Newcombe, Landau, and Maddy) pointed me toward to the theory of core number cognition as a useful framework for asking metaphysical, semantic, and epistemological questions about mathematics. The more recent literature on core cognitive theory of geometric concept acquisition seemed to offer parallel hopes for understanding geometric knowledge. In the Spring of 2016, Pen organized a reading group that allowed me to explore and discuss the claims core geometric cognition. The disanalogies I discovered during this reading group gave rise to Chapter 2.

The third and final chapter on mathematical entitlement has a longer history. In my first year at UC Irvine, after reading Giaquinto's book (mentioned above), I started discussing the philosophical significance of the nineteenth-century "flight from intuition" with Sean and Jeremy. Initially, I was searching for alternatives to set-theoretic reductionism in the philosophy of geometry, but this metaphysical concern was eventually overshadowed by my growing interest in the visual imagery content that Giaquinto appeals to in his theory of geometric knowledge. While trying to make this content logically precise, Sean Walsh introduced me to the model-theoretic field of o-minimality. After reading some of the early texts of this field, it became clear that many of these logicians developed o-minimality with some intention to make precise the spatial content of human mental representation. The content and restrictions of o-minimal structures even seemed to complement the cognitive science models of visual imagery that Giaquinto was using in his theory of mathematical entitlement. These elements combined to form the main concern of my third chapter: delimiting the scope of mathematical theorems that can be known through imagery-based mathematical entitlement.

In addition to the philosophers mentioned above, I would not have been able to finish writing this dissertation without the loving support of my family (Elliott, Harold, Jean, Karen, Mark), my partner (Elaine), and my closest friends (Sam, Alysha, Valerie, and Tyler).

CURRICULUM VITAE

J. Ethan Galebach

EDUCATION

Doctor of Philosophy in Philosophy **2018**
University of California, Irvine *Irvine, CA*

Masters of Arts in Philosophy **2015**
University of California, Irvine *Irvine, CA*

Bachelor of Science in Physics and Philosophy **2012**
University of Pittsburgh *Pittsburgh, PA*

TEACHING EXPERIENCE

Instructor (PHIL 110) **2018**
Santa Ana College *Santa Ana, CA*

Instructor (LPS 29, 60, 105, 205) **2015–2018**
University of California, Irvine *Irvine, CA*

Teaching Assistant **2012–2015**
University of California, Irvine *Irvine, CA*

ABSTRACT OF THE DISSERTATION

Representing Geometry: Perception, Concepts, and Knowledge

By

J. Ethan Galebach

Doctor of Philosophy in Philosophy

University of California, Irvine, 2018

Professor Jeremy Heis, Chair

In this dissertation, I investigate how humans represent space and other geometric entities. The topics of my three chapters are delimited by three kinds of spatial representation: perception, conception, and propositional knowledge. In chapter 1, argue against the widely-held philosophical view that the content of visual perception includes a "geometry," or more precisely, a metric space. I appeal to behavioral and neural evidence to argue that the spatial contents of visual perception exhibit a disunity that most philosophers since Kant would find surprising. In chapter 2, I evaluate a theory of geometric concept acquisition known as "core geometric cognition," and I use behavioral and neural evidence to argue that the most general assumptions of this approach are likely false. In chapter 3, I defend the view that, despite the geometric disunity exhibited by perception, it is still plausible to believe that some propositional knowledge of advanced mathematical theorems is grounded in visual imagination.

Introduction

In this dissertation, I investigate how humans represent space and other geometric entities. The topics of my three chapters are delimited by three kinds of spatial representation: perception, conception, and propositional knowledge. In chapter 1, argue against the widely-held philosophical view that the content of visual perception includes a geometry, or more precisely, a metric space. I appeal to behavioral and neural evidence to argue that the spatial contents of visual perception exhibit a disunity that most philosophers since Kant would find surprising. In chapter 2, I evaluate a theory of geometric concept acquisition known as core geometric cognition, and I use behavioral and neural evidence to argue that the most general assumptions of this approach are likely false. In chapter 3, I defend the view that, despite the geometric disunity exhibited by perception, it is still plausible to believe that some propositional knowledge of advanced mathematical theorems is grounded in visual imagination.

My first chapter begins with a demarcation of perception based on perceptual constancy that can be found in David Marr, JJ Gibson, and Tyler Burge. If we assume that the notion of perceptual constancy is capable of operationally defining *perceptual representation*, is there any reason to believe that the spatial content of visual perception includes a metric space? Analytic philosophers since the publication of Peter Strawson's Individuals (1959) have either taken for granted that the answer is 'Yes,' or have provided faulty arguments for the existence of a visuo-perceptual metric space (VMS). I use behavioral, fMRI, and EEG studies by Nancy Kanwisher, Ruth Rosenholtz, and other vision scientists to argue against the existence of a VMS. This evidence suggests that our attributions of shape to objects, agents, faces, scenes, and surfaces, all fail to meet the definition of a metric space. Why is this important? For one, it is often thought that Euclidean intuitions about parallelity, angular addition, object rotation, size comparison, etc. are embedded in the content of visual perception. Once we give up on the notion of a VMS, what explains the origins of these intuitions and what can be said of their epistemic status? Are there unlearned cognitive structures that underlie our intersubjective agreement in elementary geometry? Empirically, how intersubjective *is* this agreement? Such questions bring me to my second chapter.

My second chapter evaluation one theory of the development of Euclidean intuitions and geometric concepts known as the core cognitive theory of geometry. Philosophers of arithmetic have fruitfully drawn upon the cognitive theory of *number* to develop their metaphysical, semantic, and epistemological claims. Might a similar strategy be useful for philosophers of geometry? As the theory currently stands, this is not likely. There are a number of problems that are unique core cognitive theory of *geometry* that will need to be worked out before philosophers can rely on its theoretical claims. In particular, there is little evidence that the two core systems of geometry have the representational content and neural basis that the theory attributes to them. Moreover, the theory fails to precisely demarcate (at both an experimental and theoretical level) the concepts and beliefs it wants to explain. Finally, the theory offers much fewer details about the stages of the acquisition process for geometry than it does for arithmetic. In time, something like the core cognitive theory may end up being correct. However, the current state of scientific knowledge does not provide a firm foundation for developing claims in the philosophy of geometry.

In my third chapter, I propose an account of mathematical entitlement that is inspired by Marcus Giaquinto's book *Visual Thinking in Mathematics*. Over the past century, the philosophical literature on the epistemology of mathematics has focused primarily on axiomatic proof, sometimes even suggesting that beliefs formed outside of axiomatic deduction cannot qualify as mathematical knowledge (e.g., the Intermediate Value Theorem). As an antidote to this preoccupation with proof-based knowledge, Giaquinto's book describes how the contents of visual imagination can reliably give rise to warranted basic beliefs in the domain of elementary geometry and elementary arithmetic. In my chapter, I update and expand upon the representational contents that Giaquinto attributes to visual imagination. Additionally, I introduce the notion of a *valid connection* between imagery contents and the resulting belief as a novel means for evaluating the epistemic status of the resulting belief. Finally, I use this epistemic framework to show that both elementary and advanced formulations of the Intermediate Value Theorem are subject to imagery-based mathematical entitlement. The advanced formulation comes from a relatively new mathematical discipline known as o-minimality. It is of philosophical interest that the founders of this discipline are attempting to give an alternative foundation for geometric topology. Following Grothendieck, they insist that the objects of o-minimality are inherently visualizable and that its theorems conform to the topological intuition of shape. These assertions give additional support to my thesis that advanced mathematical beliefs can be epistemically grounded in visual imagination. As our scientific understanding of visual imagination progresses, my account will have to be updated with more accurate details about its representational contents and belief fixation processes. Despite this, I hope my epistemological account can contribute to our understanding of the nature of mathematical knowledge outside of axiomatic proof.

The unifying topic of my dissertation – spatial representation – has given me the chance to put a variety of disciplinary perspectives in conversation with one another. At times, it seemed that translating concepts and claims between these disciplines was impossible. These obstacles were temporary. I remain strongly committed to the view that any adequate understanding of how humans represent space requires us to articulate and synthesize robust normative concepts, fruitful experimental paradigms, and a large body of neural and behavioral results. No academic discipline is currently fit to do this work alone. Even if my chapters are revealed to have significant blind spots, I hope they can motivate more gifted philosophers to engage in this interdisciplinary effort to understand geometric representation.

Chapter 1

Does Visual Perception Have a Geometry?

Abstract: Over the past century, philosophers have often assumed that the content of visual perception exhibits some sort of geometric organization. During this time, there has been broad consensus that the representational contents of ordinary visuo-perceptual states include, at the very least, a metric space in which the distance relations between many visible elements are explicitly represented. This geometrically-organized representation goes by various names – "scenario content," "visual space," "reference frame," "egocentric space," "depth map," "2.5-D sketch," "coordinate space," etc. – but in this essay, I simply call it a visuo-perceptual metric space (VMS). I argue that, despite this consensus, no compelling reason has been given for believing that visuo-perceptual states contain a VMS. I also show that many neural and behavioral anomalies exist for VMS proponents. Finally, I describe our best current vision-scientific theories of the spatial contents of perception and show how they explain shape and scene perception without positing a VMS.

1.1 Introduction

Philosophers since at least the publication of Peter Strawson's *Individuals* (1959), but arguably since Kant's first *Critique* (1781), have agreed that the representational contents of visual perception are often perceptually organized in a geometric structure. For present-day philosophers that write on the spatial contents of perception, this claim – that visual perception contains a metric space – has gone largely unquestioned. When arguments for this view are put forward, they are often short and end with an appeal to allegedly supportive research in perceptual psychology.

One illustrative example of how philosophers talk about the idea that visual perception has a geometric structure comes from *The Oxford Handbook of Philosophy of Perception* (2015). Jérôme Dokic begins his essay, "Perception and Space," by declaring that the contents of perception typically have a geometric structure:

> There is a minimal, formal sense in which the perceptual field is a kind of space. The perceptual field imposes upon objects and positions a set of relations characteristic of space, and this set of relations corresponds to a mathematical structure that defines a perceptual space. ... I would like to relate these [facts] to the notion of a *frame of reference.* (2015, p 441)

By the end of his essay, Dokic has concluded that every theory of perceptual representation "should acknowledge that spatial perception ... necessarily involves frames of reference ... with different origins and coordinate systems" (p 456).[1] Like many other philosophical

[1] My concern in this essay is with the broad notion of a metric space. By positing a coordinate system, Dokic posits additional geometric structure over and above that needed for a metric space. Indeed, most philosophical proponents posit some additional structure (e.g., axes, origin, angles). However, I will largely ignore this additional structure in my argument here in order to be as expansive as possible. Still, it is important to appreciate the fact that positing a coordinate space has given rise to a number of philosophical debates about whether the geometry of perception is discrete or continuous, allocentric or egocentric, carte-

essays, Dokic largely assumes the existence of a *visuo-perceptual metric space* (henceforth, a VMS) and believes that this has been established by psychological experimentation and past philosophical argumentation.[2]

In this essay, I would like to seriously question this geometric view of perceptual representation. I believe that recent neural and behavioral evidence has revealed a deeply counterintuitive disunity in the way we perceive spatial aspects of our environment. I think philosophers should take this disunity of spatial perception seriously. Although I will not explore them here, there are significant consequences for a number of prominent debates within the philosophy of perception, such as the pictorial vs. propositional content debate and the cognitive penetrability debate.

I will argue against the existence of a VMS in three stages. First, I will show that, since Strawson (1959), the central philosophical arguments for the existence of a VMS are faulty.[3] Second, I will describe a number of experimental anomalies that *prima facie* conflict with the claim that a VMS exists. Finally, I will describe the dominant understanding of spatial representation in vision science today, taking note of which representational resources it posits in place of a VMS. As we will see, the spatial contents of perception are not organized geometrically in any robust sense.[4]

---

sian or polar, etc. These debates all assume the existence of a visuo-perceptual metric space, which reveals how entrenched the philosophical commitment to the existence of this entity has become.

[2]Dokic appeals to philosophical argumentation by Gareth Evans, a student of Peter Strawson, in his *Varieties of Reference* (1982).

[3]Although he limits his focus to *egocentric* coordinate spaces, David Bennett (2016; cf. 2012, p 25) is the only philosopher that I am aware of who questions the existence of a VMS. He offers a critique of some of the more recent philosophical arguments for the egocentric VMS. The arguments that he rebuts – arguments based on visual processing of surface curvature, structure-from-motion, etc. – do not overlap with the more historically influential arguments that I consider here. I am in agreement with many of Bennett's conclusions.

[4]Perceptual states may represent spatial properties – shape, size, order, containment, distance, direction, angle, length, orientation, etc. – properties without *thereby* representing a metric space. Do our spatial representations achieve a geometric unity in post-perceptual processing? This question is partially addressed in chapters 2 and 3.

1.2 VMS Proponents and Arguments

Philosophers have used a wide variety of terms to refer to visuo-perceptual metric spaces. The most common of these terms are: *visual space, 2.5-D sketch, scenario content, egocentric space, reference frame, coordinate system, visual field,* and *depth map.*[5] Yet, each of these spaces has two features in common that I will use to define the notion of a VMS. First, they all contain a set of distance relations that hold between most, if not all, entities in the space. These distance relations are symmetric (i.e., $d(x, y) = d(y, x)$) and obey the triangle inequality (i.e., $d(x, y) + d(y, z) < d(x, z)$). Second, they are all part of the representational content of a visuo-perceptual state. Before I consider the specific philosophical arguments offered for the existence of a VMS, a few caveats are in order regarding the meaning of "visuo-perceptual state."

1.2.1 What is Visual Perception?

As I will be using these terms, *perception* is importantly different from *sensation* and *cognition*. I will start with the sensation/perception border. The main feature that distinguishes sensory states from perceptual (and cognitive) states is the fact that they are not *representations*. That is, sensory states lack *accuracy conditions*. Sensory states necessarily lack representational content, while perceptual states necessarily have content.[6] I follow philosophers like Tyler Burge (2010, pp 92), and perceptual psychologists like J. J. Gibson (1966, p 1), in maintaining that the most primitive type of representational content arises via *con-*

[5]For "visual space," see Hatfield (2009, p 5). For "2.5-D sketch," see Hatfield (2009, p 5), Green (2015, p 5). For "scenario content," see Peacocke (1992, pp 64ff), Heck (2007, p 14), Schellenberg (2008, p 61). For "egocentric space," see Evans (1982, p 163). For "reference frame," see Burge (2010, pp 199-201), Dokic (2015, pp 443ff); Campbell (1994, ch. 1). For "visual field," see Strawson (1959, p 65). For "coordinate system," see Wu (2013, p 652), Bermudez (1998, p 141). For "depth map," see Bennett (2016, p 6).

[6]I will use "representational content" and "content" interchangeably. Some philosophers, like Gary Hatfield and Christopher Peacocke, describe the phenomenal character of an experience as part of its "content," but I will not do so here.

stancy in cortical visual areas. Constancy transformations issue mental states with accuracy conditions involving perspective-invariant distal properties.[7] These mental states represent, e.g., the *shape* and *color* properties of distal entities. On this understanding of perception, the retinal image – i.e., the sensory registration of luminance on the retina – does not constitute a visuo-perceptual state. Hence, the retinal image is not a VMS despite the fact that it can be usefully analyzed as a metric space. My thesis is not so bold as to claim that the retinal image does not exist.

The perception/cognition border is more controversial among philosophers. There are two schools of thought on how to draw this distinction.[8] The first school of philosophers demarcates perception from cognition in terms of an allegedly distinct type of *phenomenal character* that perceptual states have. On this understanding, a perceptual state can normally be introspectively individuated and classified as perceptual by the person in that state.[9] Philosophers in the second school find this phenomenological demarcation of perceptual states to be "deeply wrong-headed" (Burge 2014, p 11). This second school rejects the claim that introspection can determine whether a given mental state is perceptual or cognitive:

[7]This has been a common restriction since at least Gibson (1966, p 1) and Marr (1982, pp 29, 34, 36), which call these perspective-invariant properties "constant," "valid," "invariant," "objective," "physical," and "permanent." Marr distinguishes 'primitive' visual systems, like that of a fly, from 'advanced' visual systems, like that of a human, on the basis of the latters ability to attribute perspective-invariant properties like 3D-shape (pp 32-36). Similar appeals to perspective-invariance as constitutive of constancy can be found at Palmer (1999, p 125), Nakayama and Shimojo (1992, p 1357), and Bennett (2016, p 4). Today, determinations of perceptual constancy are most often operationalized in fMRI adaptation effects because "fMRI ... adaptation across two different stimuli provides evidence for a common neural representation" (Epstein, et al. 2017, p 1505; cf. Kanwisher and Dilks 2014, p. 733). Behavioral paradigms are also used to determine constancies: aftereffects (Marr 1982, p 287), dishabituation effects (Kayaert and Wagemans 2010), categorization, individuation, and reidentification abilities (Potter 1975, 2014; Sekuler and Palmer 1992; Chen 1982, 1990, 2005; Rosenholtz 2016), tracking abilities (Zhou et al 2010), and illusions (Chen and Zhou 1997, Zhou et al. 2003, Todd and Bressan 1990).

[8]Orlandi (2014) mentions a similar distinction between two schools of philosophy of perception: those who reflect on "conscious perceptual experience" and those who reflect on the "current best models of perception in the cognitive sciences" (p 13). See also Nanay (2017), pp 2-3.

[9]See Peacocke (1992, p 62), and Siegel (2010, pp 3-4). As Siegel (2017b) puts it, "We have an understanding of perceptual experiences that comes from our familiarity with them, and that understanding is robust enough to identify the experiences, but not detailed enough to settle the [question of which properties are perceptually represented in a given experience]" (p 74).

> Can we be sure from introspection that [the contents and phenomenal character of an experiential state] are really perceptual, as opposed to primarily the "cognitive phenomenology" of a conceptual overlay on perception, that is, partly or wholly a matter of a conscious episode of perceptual judgment rather than pure perception? (Block 2014, p 7)

Philosophers like Block and Burge deny that perceivers have introspective access to their own mental states in a way that allows them to individuate or classify these states as perceptual or cognitive. Instead, they offer a perception/cognition demarcation derived from vision science. According to common usage in vision science, a visuo-perceptual state necessarily (a) finds its neural correlate in the ventral or dorsal visual pathways,[10] and (b) is the result of processing information from a single fixation.[11] If a representational mental state definitively does not meet one of these criteria, it is post-perceptual and hence *cognitive.*

These two perception/cognition demarcations, one phenomenal and one vision-scientific, issue different verdicts about whether perceptual states can represent, for instance, a laptop, a punt return touchdown, your own neighborhood, or that two complex three-

[10]Most vision science focuses on the belief-guiding content of the ventral pathway, which goes from the occipital to the temporal lobe and includes visual areas like LOC, PPA, OPA, EBA, FFA, and VWFA (Kanwisher and Dilks 2014). Less is known about the action-guiding content of the dorsal pathway(s), which goes from the occipital to the parietal lobe (Kravitz *et al.* 2011; Konen *et al.* 2013; Bell *et al.* 2014).

[11]This restriction has been common since at least Julesz (1975), which distinguishes single-fixation "pre-attentive perceptions" of texture properties from transsaccadic "scrutiny" representations. A single fixation lasts for "300 msec on average" (Henderson and Hollingworth 2003, p 58). Psychologists who explicitly endorse this single-fixation view of perception include: O'Regan (1992), Rensink (2000, 2002), Henderson and Hollingsworth (2003), and Fei-Fei *et al.* (2007). Some smaller (and largely independent) traditions of research in psychology departments – namely, visual psychophysics and computer vision – do not clearly distinguish perception from sensation or cognition, and hence I will not include their theoretical paradigms within the category of vision science *pace* Palmer (1999).

dimensional shapes are identical.[12] In this essay, I will follow the vision-scientific demarcation in specifying the scope of the VMS-concept.[13]

Now that we have a better sense of the scope of the VMS-concept, I will lay out and critique the views of five prominent philosophers who endorse a VMS: Peter Strawson, Christopher Peacocke, Tyler Burge, Gary Hatfield, Wayne Wu.[14] These five figures are not the only proponents of a VMS, but they are the only ones I have found that give arguments for believing that a VMS exists.[15]

1.2.2 Five Arguments for a VMS

The idea of a visuo-perceptual metric space is not new. In the first *Critique*, Kant posits a sensation-involving nonconceptual representation of space, *intuition*, that at the very least obeys the triangle inequality:

[12] In this footnote, I say why these four items cannot be represented perceptually if we accept (a) and (b). Kind-attributions like *this laptop* will have neural correlates outside of early visual areas (Kim *et al.* 2009). Event-attributions like *this punt return touchdown* will be based on information from dozens of saccades and fixations in various memory stores. Scene reidentification like *this is my neighborhood* (from Siegel 2017, p xiii) will likely involve activating cognitive maps from long-term memory, which seems to take place in RSC, an area whose perceptual status is unclear (Epstein and Vass 2013). Finally, shape-identity judgments like *these two objects have the same shape* (from Peacocke 1992, p 78) will result from mental rotations in visual working memory. These four representations will all likely count as perceptual under the phenomenological demarcation, but they are likely post-perceptual under the vision-scientific demarcation.

[13] I do this because I agree with Block and Burge that our introspective abilities are not strong enough to make this distinction in many cases. Hence, what vision scientists find theoretically fruitful will be a preferable distinction to the phenomenal one. Moreover, I have sympathies with the 'reductive' approach to phenomenal consciousness known under the label "higher-order theories of consciousness" (Carruthers 2016).

[14] Two philosophers that I will be discussing in detail, namely Peacocke and Hatfield, do not make clear efforts to distinguish between these two demarcations of perception from cognition. As such, it is logically possible that they would reject the existence of a VMS as I have narrowly defined it. However, since both philosophers substantively engage with vision-scientific literature, they may endorse it. I will proceed as if Peacocke's scenarios and Hatfield's visual spaces are both VMSs. Even if this is assumption is inaccurate, their arguments for the existence of scenarios and visual spaces are unconvincing as I will show.

[15] Other twenty-first century proponents include Richard Heck (2007, p 14), David Chalmers (2006, p 22), Mohan Matthen (2005, p 300), Susanna Schellenberg (2008, p 61), Robert Briscoe and John Schwenkler (2015, p 1458), and Brad Thompson (2010, p 140).

> Space is not a discursive, or as one says, general concept of relations of things in general, but a pure intuition ... [G]eometrical propositions, that, for instance, in a triangle two sides together are greater than the third, can never be derived from the general concepts of line and triangle, but only from intuition and indeed *a priori* with apodictic certainty. (Kant 1781, A24-5/B39-40)

Throughout the twentieth century, dozens of philosophers have elaborated this metric-space conception of perceptual content.[16] I will focus on what I regard as the five most substantive arguments for a VMS, namely, the arguments of Strawson, Peacocke, Hatfield, Burge, and Wu. I will address each of these figures in turn, first demonstrating that they posit a VMS and then laying out their argument. Each argument claims that humans would be unable to complete a certain visual or cognitive task without a VMS. After laying out each argument, I will show why it is flawed.

Strawson

In his 1959 book *Individuals*, Peter Strawson endorses a VMS in the following passages. He claims that "the visual field is necessarily extended at any moment, and its parts must exhibit spatial relations to each other" (p 65). The locations of these *spatial parts* (also called *visual elements*) are perceptually specified in terms of "a common reference point and common axes of spatial direction [and distance]" (p 22). According to Strawson, these "visual elements can be seen all at once as at a certain visual distance from one another ... Or, to put it in another way: the momentary states of the colour-patches of the visual scene visibly exhibit spatial relations to each other at a moment" (80). From these passages, and

[16] The claim the visual perception has a geometry akin to a VMS can be found in the following twentieth-century texts: Mach (1902), Carnap (1922), Cassirer (1944), Strawson (1959), Grünbaum (1962, 1973), Putnam (1963), Sellars (1968, pp 25-30), Suppes (1977, 1995), Taylor (1978), O'Shaughnessy (1980, p 176), Evans (1982, ch 6), French (1987, p 115), Baldwin (1992, p 183), Peacocke (1992, ch 3).

many others like them, it is clear that Strawson conceives of visual perception as containing a (coordinatized) metric space.

What is Strawson's argument that such an entity exists? He claims that an adequate conceptual analysis of our capacity for linguistic reference requires the existence of a VMS. His argument rests on a few premises, the first of which is the claim that every instance of linguistic reference is also an instance of *identifying* reference (p 16). According to Strawson, a speaker makes an *identifying reference* only if she knows an "individuating fact" about the referent (p 23). So, an implication of Strawson's first premise is that linguistic reference to object a, achieved by uttering an expression E, requires, for some concept P, that the speaker knows: E refers to a & $P(a)$ & $\forall y(P(y) \rightarrow a = y)$. The proposition, $P(a)$, is the individuating fact by which the expression E gains its "demonstrative force" (p 118).

Strawson's second premise is that the concept P must individuate the referent by (i) conceptually locating it within a region of a spatio-temporal coordinate space (pp 38, 56), and (ii) providing a description that is true of the referent and false of the other particulars in that region (p 25). Why does Strawson privilege the spatio-temporal coordinate space as the means of all identifying reference? He does this because such a space is "uniquely efficient" at generating the individuating facts required for identifying reference (p 24).

The third premise of Strawson's argument is that the "visual field [must provide] the materials for spatial concepts" (p 65).[17] By this claim, Strawson means that the visual *perception* must have the same geometric structure as the spatio-temporal coordinate space underlying the individuating concept P. Hence, Strawson conceives of his "visual field" as necessarily a visuo-perceptual metric space (p 65).

Taken together, Strawson's three premises entail that the cognitive task of linguistic reference requires a VMS. But his argument is not sound. Strawson's first and third premises

[17]Strawson considers but rejects the view that auditory perception would also "suffice to generate spatial concepts" (p 66).

are false. The first premise is false because linguistic reference does not require knowledge of an individuating fact. This is now orthodoxy in philosophy of language and is associated with the general view known as "semantic externalism." Most philosophers agree that certain examples of everyday reference made famous by Kripke (1972) demonstrate that a speaker need not know any individuating facts about a referent in order to refer to it. Strawson is simply wrong when he claims that "one cannot significantly use a name to refer to someone or something unless one [is] prepared to substitute a description for the name" (p 181).

A neo-Strawsonian might try to salvage the argument from reference by granting that semantic externalism holds for linguistic reference, but still claiming that *perceptual* reference requires attribution of an individuating location to the referent. Indeed, this is precisely the view of Gareth Evans. Evans embraces semantic externalism in his 1973 essay, "The Causal Theory of Names," but Evans later claims that *perceptual* reference, unlike reference from memory or testimony, requires "a conception of it as the occupant of such-and-such a position" in an egocentric frame of reference (1982, p 149).[18] Evans argues that this latter claim is a consequence of his *Generality Constraint*, which Evans states as: "If we hold that the subject's understanding of 'Fa' and his understanding of 'Gb' are structured, we are committed to the view that the subject will also be able to understand the sentences 'Fb' and 'Ga' " (p 102). This constraint allegedly entails that perceptual reference is only possible when a perceiver is capable of *reidentifying* the object (p 149). And reidentification is allegedly accomplished on the basis of the object's perceived location within an egocentric frame of reference (p 149).

There are a number of problems with Evans' argument. For one, it is not clear why Evans requires *locative facts* to be the basis of reidentification rather than, say, *shape facts*. For another, it seems fairly clear that we are capable of individuating, tracking, reidentify-

[18]Cf. Evans (1985), p 392. More recently, Evans' location-based view of perceptual reference has been endorsed by Robert Briscoe: "To see a matchbox as over there, e.g., is perforce to see it as located somewhere relative to here, somewhere, that is, more precisely specified using the axes *right/left*, *front/behind*, and *above/below*" (2009, p 424).

ing, and thinking about objects that we unwittingly see in a mirror. Such objects are not perceived in their true location and so Evans is committed to the unappealing view that this is a case of reference failure.[19] Finally, the Generality Constraint is primarily about *thoughts*, not *perceptual representations*. Even if we extend it to perception, Evans does not say how his reidentification requirement could be derived from the Generality Constraint.[20] Hence, a neo-Strawsonian view of reference is without any argumentative support. Moreover, many cogent philosophical arguments for semantic externalism regarding perception have been developed, although I will not review them here.[21]

Strawson's third premise is also false. It is not true that the geometric character of perception must be the same as the geometric character of our conception of the world. Even if we grant to Strawson the bold claim that humans conceive of all objects as located in a single, unified coordinate space (p 31), why should we think such a metric space is present in perception? I assume Strawson would say that we could not develop (or "fill in") our *conception* of the world without a VMS. But this is simply inaccurate. It is logically possible that each perceptual state represents a single distance relation between two objects, and only later in a post-perceptual region (e.g., the hippocampus) do we integrate these relations into a unified coordinate space.[22] It is even possible that we develop (and "fill in") our conception of the world without incorporating *any* perceived metric properties of our environment – we

[19] I draw this objection from Tyler Burge (2010, p 200) and E. J. Green (2017, p 14). Burge and Green each provide some other cogent objections to Evans' locative requirement on perceptual reference. For instance, Burge says: "One can see and think about a star or comet through light that is refracted by the atmosphere. One might have seriously mislocated the object with respect to ones own position and have no practical way of locating it correctly" (p 200).

[20] It seems very likely that Evans is drawing his reidentification requirement from Strawson (1959, ch. 2). But Strawson's reidentification requirement is (a) a requirement for *linguistic* reference, (b) grounded in the capacity for sortal predication, and (c) lacking any argument for why every sortal predication contains the idea "of a continuous path traced through space and time" (p 207), let alone for Evans' further implication that the perception of *path-continuity* requires the perception of a coordinatized metric space.

[21] Most prominently, Burge's defense of "anti-individualism regarding perception" (1991; 2010, ch. 3).

[22] This might be how the hippocampus constructs a cognitive map representation of one's environment, but it is still unknown which perceptual and cognitive contents are integrated during this construction process. For an overview of the currently-available neural and behavioral evidence on this topic, see Epstein *et al.* (2017).

might use saccadic corollary discharges, not perceptual contents, as our basis for adding new objects and distance relations to our conception of the world cognitive map.[23]

In sum, Strawson's argument from linguistic reference fails to provide any support for the claim that a VMS exists because his first and third premises are false. Although Strawson's argument is not often explicitly endorsed today, his influence is still felt through Evans and one of Evans' colleagues: Christopher Peacocke. Peacocke provides our second argument for a VMS.

Peacocke

In his 1992 book *A Study of Concepts*, Christopher Peacocke posits a VMS he calls a "scenario."[24] According to Peacocke, a scenario is a perceptual representation that specifies the coordinate locations and properties of all the surfaces in ones field of vision.[25] Peacocke claims that scenarios are perceptually ascribed to the unique physical location "given by the property of being the center of the chest of the human body, with three axes given by the directions back/front, left/right, and up/down with respect to that center" (p 62). For Peacocke, the representational content of a scenario specifies,

> for each [discernible] distance and direction from the origin..., whether there is a surface there and, if so, what texture, hue, saturation, ...brightness, ...degree of solidity, ...orientation, ...it has at that point. (p 63)

[23]Cf. Redish (1999), pp 84, 85, 277. Corollary discharges are visuo-motor neural signals that 'predict' future retinal stimulation after a saccade on the basis of a saccadic motor command.

[24]Peacocke's postulation of scenarios was directly inspired by Gareth Evans' above argument for a VMS. See Peacocke (1992, p 71) where he endorses "Evans' Thesis."

[25]Other VMS proponents speak not of component *surfaces* but component points, textures, bars, blobs, luminance gratings, contours, volumetric objects, enclosures, etc. I will gloss over these differences about the elements that constitute the underlying set of the posited metric space.

These contents determine the accuracy conditions of a scenario representation. When the *environmental* surfaces match the *represented* surfaces with respect to the features mentioned in this quote, the perceptual state with the scenario content is considered to be accurate.[26]

Peacocke's initial motivating example for believing in a VMS is phenomenological. He claims that distinctions in the way visual experiences *feel* should be explained in terms of their distinct perceptual content:

> The appropriate set of labeled axes captures distinctions in the phenomenology of experience itself. Looking straight ahead at Buckingham Palace is one experience. It is another to look at the palace with ones body turned toward a point on the right. In this second case the palace is experienced as being off to to one side from the direction straight ahead, even if the view remains exactly the same as in the first case. (p 62)

In other words, Peacocke claims the distinct feel of these two experiences of Buckingham Palace supports the claim that a surface's direction from the perceiver's body is perceptually represented.[27] In doing so, Peacocke is assuming that phenomenal difference always supervenes on representational difference. This view is usually called *representationalism*.[28] However, in his earlier book, *Sense and Content*, Peacocke himself seems to reject representationalism when he claims that the two phenomenally distinct interpretations of the Necker cube have the same perceptual content (1983, pp 16-17). Representationalism is also

[26] Do scenarios qualify as *visuo-perceptual* in my sense? It seems so. Peacocke says that scenarios are "most primitive level of ... nonconceptual representational content" (90). This suggests that scenarios are a part of the content of visuo-perceptual states as defined by (a) and (b) above. Moreover, Peacocke discusses and partially endorses the representational theories of early vision given by vision scientists like David Marr, Stephen Palmer, Anne Treisman, and Roger Shepard. These vision scientists are explicitly concerned with early visual areas, perceptual constancies, and visual representations derived from single eye fixations. So, although Peacocke accepts the phenomenal demarcation of perception from cognition, his scenarios seem to be perceptual in my sense.

[27] Peacocke formalizes this directional relation as, "R is located in direction D" (p 70). Presumably, this direction D is encoded as three numbers representing the degrees of each angle formed by the vector to the surface and each of the body's three axes.

[28] Cf. Lycan (2015).

fundamentally at odds with the vision-scientific definition of perceptual content in terms of perspective-invariant constancies.[29] Hence, this initial motivating example does not constitute much of an argument.

Later on, Peacocke gives two additional reasons for believing in a VMS. First, he claims that if "you are looking at a range of mountains, it may be correct to say that you see some as *rounded*, some as *jagged*. But the content of your visual experience in respect of the shape of the mountains is far more specific than that description indicates" (p 67). Peacocke is claiming that when we perceive a scene like a range of mountains, the content is more pictorial than conceptual, and that pictorial content should be articulated as a VMS. The problem with this argument is that almost every current vision-scientific model of scene perception understands "pictorial content" in terms of the global texture properties of the retinal image rather than as a scenario representation. There is good experimental and neural evidence for this view as we will see in section 3. One of the more revealing experimental results from Mary Potter and colleagues demonstrates that, even though we can *categorize* a scene as, e.g., 'mountains' or 'flowers,' after only seeing it for a few dozen milliseconds, we are unable to reidentify *which* flower scene we saw even when the alternative choice has a vastly different metric layout (Potter 1993; Potter *et al.* 2014). Over longer viewing times, when we can distinguish *round* from *jagged* mountains, it is likely that the underlying perceptual content is given by *texture* properties rather than a VMS. Hence, Peacocke's first argument is weak.

Peacocke's second argument amounts to the claim that scenario content is the basis of our capacity to "confirm or refute" the accuracy of our perception of an object's shape

[29] See footnote 7. This is not to say that perceptually representing body-centric direction is impossible – such a representation might be invariant over eye or head motion – but Peacocke's Buckingham Palace example does nothing to make this possibility more likely. Moreover, even if Peacocke is right that we perceive the body-centric direction to Buckingham Palace, there are many ways to articulate this content without positing a VMS. For instance, some visual neuroscientists and developmental psychologists have claimed that we perceptually categorize an object as *to the left* or *to the right* independent of any metric directional content (Scott *et al.* 2016; Gava *et al.* 2009).

(p 242n11). This raises the question: If I perceive that *this cube is red*, can I refute this perception *without* perceptually locating the cube in a VMS? This seems eminently plausible. In order to refute this perception, I will need to reidentify the cube, and this can be done in a number of ways without a VMS. For example, I might reidentify the cube on the basis of some rare markings or the object next to it. There are many ways to reidentify an object without using a VMS. Hence, Peacocke's arguments do not provide any support for believing in a VMS.

Hatfield

Gary Hatfield is our third VMS proponent. He describes his VMS in his theory of "visual space." Hatfield claims that visual space is "like David Marr's 2.5-D sketch" (2009, p 5).[30] David Marr's 2.5-D sketch will be discussed more fully in section 3, but here I will simply note that it contains a three-dimensional coordinate space centered on the ego. Presumably Hatfield's visual space also has this property. Regarding its metric structure, Hatfield claims "that a finite Euclidean model captures the gross structure" of visual space (2003, p 172). Visual space, like Marr's 2.5-D sketch, is supposed to "guide action and ... support reasonably accurate judgments of size, distance, and shape" (2009, p 205). So it seems that visual space arises early in the visual cortex before processing splits into action-guiding dorsal content and judgment-guiding ventral content.[31] Nevertheless, there are two notable differences between Hatfield and our previous proponents. First, Hatfield thinks of visual space in the first instance as part of the phenomenal character of a perceptual state. Hence, we determine the three-dimensional geometric structure of a visual space by asking perceivers

[30]The psychologist Rudolf Luneburg (1947) introduced the idea of a "visual space" with a hyperbolic metric as a theory of the perceptual content underlying a particular visual task: judging the parallelity of black cords or a series of lights. The "visual space" continues to be a topic of investigation in the (to my my mind rather speculative) discipline of mathematical psychology (Indow 2004, Wagner 2005, Koenderink and Van Doorn 2008), and in most of these cases visual space has an egocentric coordinate system

[31]See footnote 10.

how things *appear* to them.[32] Nevertheless, Hatfield thinks the geometric structure of visual space *determines* the perceptual state's geometric representational content via a Euclidean transformation (2013, p 59).[33] For example, even if two parallel railroad tracks converge in (phenomenal) visual space, Hatfield thinks we often perceptually *represent* the railroad tracks as parallel. As long as the phenomenal visual space successfully "serve[s] action guiding, object cognition, and other purposes," Hatfield thinks our perceptual state can be said to accurately represent the geometry of the visible environment (2013, p 59). So it should be no surprise that Hatfield's primary argument for a VMS is based on psychological experiments in which participants estimate various metric properties of their environment. Hatfield thinks that these experimental results – particularly the object size estimation results of Granrud (2004, 2013) – support the view that children develop a veridical VMS by the age of 10 (2013, p 49).

Hatfield's argument for a VMS has many shortcomings. Most significantly, it assumes that our perceptual attribution of an object's size in embedded in a unified geometric representation of the whole visual environment. It is of course true that we *feel* that an object's size is perceptually given to us in a geometrically organized environment. But this *feeling* substitutes the phenomenal definition of perception for our preferred vision-scientific definition. In visual neuroscience today, it is near orthodoxy that *object size* and *spatial layout*

[32]For Hatfield, perceptual states have two aspects: a phenomenal aspect and an representational aspect (2009, p 18). The phenomenal aspect of a perceptual state is a collection of subject-dependent entities, properties, and relations. The phenomenal aspect *does not purport to represent* anything in the visible environment. For instance, when I look at the Kanisza triangle illusion and assert *that there appears to be a bright triangle there*, Hatfield would say that this assertion reports the phenomenal aspect of my perceptual state rather than the representational aspect. Hatfield claims that the phenomenal aspect of a perceptual state has its own "spatial structure" (p 18). On the other hand, the representational aspect of a perceptual state are the source of the states veridicality conditions and it is not reducible to biological function (p 25). For Hatfield, it is fundamentally a scientific task to specify what a perceptual state represents, and ones evidence will come from behavioral, evolutionary, and neurophysiological evidence. Hatfield's two aspects of perception allow him to claim that a single perceptual state may have phenomenally present a *trapezoid* while representing a *rectangle* (p 19).

[33]Although this fact is not relevant to his argument, the Euclidean transformation is allegedly needed to explain systematic underestimations of metric properties (2009, pp 171-175, 183). Outside of the visual space framework, other theories could be developed to explain these underestimations (e.g., maybe Intraub's metric-free notion of "boundary extension" in scene representations could account for them (Intraub and Richardson 1989; Intraub 2014)).

constancies are processed by two distinct neural pathways in the ventral stream. Object size constancy is represented in the lateral occipital cortex (LOC), whereas spatial layout constancies are represented in the parahippocampal place area (PPA).[34] Hence, even if our verbal estimation of an object's size is drawn directly from a perceptual representation of the object's size (as Hatfield supposes), there is substantial neural evidence to believe this perceptual representationdoes not determine the geometry of the perceived spatial layout (i.e., "visual space").

A second, more narrow, objection to Hatfield's argument comes from the experiments in Foley (1972). Foley's experiments demonstrate that visual space, supposing it exists, does not have a constant curvature geometry (p 328). In his second experiment, Foley asked participants (standing at a location O with a fixed luminous point A in front on them) to first place a luminous point B in the right side of their visual field to create an isosceles right triangle, OBA.[35] Next, they were asked to place a luminous point C in the left side of their visual field to create an isosceles right triangle, BAC. Finally, participants were asked whether OA and BC have the same length. If a participant answered "no," they were also asked to estimate the ratio of the lengths of OA and BC. On average, he found that BC was perceived to be 20% longer than OA (p 327). The results of this study are often taken to show that, at best, the metric structure of visual space violates the "axiom of the congruence of triangles," and thus cannot have "any constant curvature geometry, including Euclidean geometry" (Masrour 2015, p 1823).[36] At worst, these results suggest that attempting to describe the unified geometric structure of visual perception is a red herring. Either way, Foley's experiments disconfirm Hatfield's claim that visual space is Euclidean. In sum,

[34]For an illustrative example, see Park *et al.* (2011). To get a fuller picture, see the essay collection *Scene Vision* edited by Kveraga and Bar (2014).

[35]More exactly, Foley asked participants to place point B so that (i) OB is perpendicular to BA, and (ii) OB and BA have the same length (1972, p 326).

[36]Suppes (1995, p 39), Wagner (2006, p 29), and Masrour (2015, p 7) all use this result to justify their rejection of a constant curvature geometry for visual space. A somewhat analogous result for *object shape* perception is presented in Domini and Braunstein (1998).

Hatfield's argument from size estimation provides no support for believing that his visual space exists.

Burge

Tyler Burge (2010, p 200) draws his conception of a VMS from Gareth Evans' *Varieties of Reference*. Burge rejects Evans' referential argument for a VMS (outlined above in the *Strawson* section), but he endorses a different argument that he attributes to Evans.[37] According to this argument, "egocentric spatial frameworks are necessary to spatial perceptual representation [because they] figure centrally in agency" (p 201). Burge main evidence for this claim is that, "empirically it is nearly certain that some animals have egocentric spatial perceptual abilities" (2010, p 207).[38] His only cited example from experimental psychology is a 1998 article by Fred Dyer on bee navigation. Burge describes this article as defending the view that "the perception and perceptual memory involved in this particular navigational task use only egocentric spatial frameworks" (2010, p 202). However, Dyer describes the bee's navigation abilities as grounded in its encoding of a "retinally localized snapshots" (p 148). When this retinal image is presented again, Dyer claims it triggers the bee to fly in a specific direction. Clearly, Dyer's conception of the bee's egocentric framework does not constitute a *perceptual* state in my (or Burge's) sense.

Burge (2014) gestures at a second argument for a VMS based on *object shape* perception. He claims that perceptually representing an object *as a cube* requires plotting straight lines in a three-dimensional coordinate system:

> Visual perception occurs in an egocentrically anchored, spatial coordinate system. ... [The] edge of a cube must be specified not merely as an edge, but through

[37] Burge (2010), pp 200-201. See also footnote 19.

[38] Burge also states this claim in an older essay when he says: "attribution of egocentric indexes is ubiquitous in perceptual and animal psychology" (2003, p 330).

> specifications that plot the edge in the coordinate system – give its length, shape, and orientation, using spatial specifications within the coordinate system. (p 492)

As before, it is not obvious that perceiving an objects shape requires perceiving its location. Indeed, Burge (2010, p 202) makes this very point in arguing against Evans.[39] On an empirical level, the neural and behavioral evidence regarding object shape perception suggests that shape attributed is size-invariant and orientation-invariant.[40] Hence, it seems that a perceptual attribution of shape does not necessitate a perceptual attribution of edge "length" or "orientation" as Burge claims. Burge might respond that this representation is computed on the basis of his proposed VMS shape representation. This is logically possible, but today the most widely adopted computational model posits that a collection of 2D images, rather than a VMS, is the basis of object shape attribution.[41] There may be a compelling argument for a VMS on the basis of object shape perception, but Burge does not provide it.

Wu

The final argument for a VMS is given by Wayne Wu. Before we look at Wu's argument, I want to establish that Wu actually posits a VMS. Consider the following passage where he describes the representational content of visual experience:

> Since normal human visual experience is three-dimensional and the locations of visible objects are represented relative to the egocenter, we can present egocen-

[39] Peacocke (1992, p 242n11) puts forward a compelling argument that shape perception does not require location perception: "Visually disoriented subjects are able to identify and apparently perceive the shape of objects in their environments without experiencing them as have any particular (egocentric) location. There is a readable case study in Godwin-Austen 1965. ... I thus allow that something is perceived as square without being localized."

[40] Konen and Kastner (2008), Kourtzi *et al.* (2003), Grill-Spector *et al.* (1999), Biederman and Cooper (1992).

[41] This is the HMAX model of Reisenhuber and Poggio (1999). For an overview of the current status of this model, see Li *et al.* (2015). Opponents of the HMAX model, like Biederman, admit that it is the "default framework" for 3D object shape perception in vision science (Hayworth *et al.* 2011, p 1).

> tric spatial content in a Cartesian coordinate system, centered on part of the perceiver, with the Cartesian axes setting egocentric directions (e.g. the z-axis defines straight ahead). More appropriate would be a spherical [i.e., polar] coordinate system that represents egocentric distance explicitly as a vector originating at the egocentric origin and whose magnitude is the distance between visible objects and the reference point. (2014, p 391)

In this passage, Wu technically claims that "visual experience," not "visual perception," contains a three-dimensional coordinate space. Is there reason to think that visual perception contains this coordinate space? I think so. For one, Wu claims that visual experience occurs in a "visual system" that is distinct from a "cognitive" system (2013, p 648). For another, he says that we locate objects in this space during a single fixation (2014, p 394). Wu provides further confirmation that his coordinate space is visuo-perceptual when he cites Peacocke's writing on scenarios as a way of articulating his own view on the "egocentric spatial content" of perception (2013, p 653n7). On this basis, I claim that Wu posits a VMS. Wu's VMS is unique in that it has polar coordinates *pace* Strawson, Peacocke, and Burge, each of whom suggest that the VMS has Cartesian coordinates. But Wu, like most philosophers since Kant, accepts that visual perception has a geometry about which we may inquire, "How is it coordinatized?"

Wu's argument for a VMS is grounded in the claim that a VMS is required to explain *position constancy* and *motion constancy.* Position constancy is the capacity to accurately perceive objects as stationary after a saccadic eye movement (despite changes in the retinal image before and after the saccade). Motion constancy is the parallel capacity to accurately perceive objects as in motion after a saccadic eye movement. Wu claims that these capacities exist and that the standard non-VMS explanation of these capacities, "the Transsaccadic Memory Account" (TM), is inadequate. His alternative explanation posits a VMS. It will be sufficient to refute Wu's argument against the TM account of position constancy and motion constancy.

According to the TM account, object motion attribution is computed purely on the basis of comparing retinal images before and after a saccade.[42] The first retinal image merely needs to be adjusted on the basis of the corollary discharge from the eye saccade. This account of motion constancy makes no appeal to a VMS since retinal images are sensations, not representations.[43] Wu (2013, p 651, fig 1) provides a helpful diagram of this allegedly inadequate account. Wu claims that the TM account falsely assumes that, "where error is high, the subject experiences visual spatial inconstancy [i.e., motion]; where error is low, the subject experiences spatial constancy [i.e., no motion]" (p 651). According to Wu, this assumption is false because

> error signals are deployed in other areas in the brain that have nothing to do with spatial constancy but where prediction is useful (e.g. in motor control, and in explaining auditory verbal hallucination in schizophrenia). The Transaccadic [*sic*] Memory account is explanatorily incomplete at a critical point. (2013, p 651; citations removed)[44]

In other words, Wu is claiming that low error signals in the above-cited diagram cannot be necessary and sufficient for perceptual motion attribution. After all, error detection is used in many non-perceptual systems. However, this is a straw man. I am not aware of any TM proponent who has claimed that *any* low error signal in *any* brain area is sufficient for perceptual constancy, let alone position constancy. Most proponents of the Transsaccadic Memory Account merely claim that *in the case of early vision* this particular error signal is

[42] Cf. Wurtz *et al.* (2011), p 496.

[43] Wu (2014, p398) attributes this Transsaccadic Memory Account to the neuroscientists who wrote Wurtz *et al.* (2011). This account is also endorsed by Henriques (1998), who defends it against various alternative accounts that posit a VMS.

[44] A virtually identical passage (and argument) is presented at Wu (2014), p 398.

necessary for the constancy transformation underlying motion attribution.[45] Until Wu states more fully his objection to the TM Account, we have no to reject this non-VMS account of position and motion constancy. That is, I do not think position or motion constancy considerations provide any grounds for positing a VMS.[46]

In sum, all of the arguments for a VMS from our five proponents are seriously flawed. However, the idea of a VMS did have an explanatory purpose. The VMS was supposed to be a single representation that explains human performance on a number of visual tasks such as motion constancy, scene categorization, shape perception, linguistic reference, location reports, and cognitive map construction.[47] In some sense, the only remaining virtue of the VMS is the unity it gives to our various explanations of these visual tasks. Hence, in section 3, I will show in some detail that current neural evidence supports a radically disunified approach to the spatial contents underlying our most common visual tasks. Before we turn to this, however, I would like to discuss two well-isolated behavioral anomalies that count against the existence of a VMS: distance reports and visual crowding. In the next section, I will show why these two visual tasks support my claim that a VMS does not exist.

1.2.3 Behavioral evidence against a VMS

Philosophical proponents of the VMS since Evans have distinguished themselves from older philosophers (like Strawson and Kant) by appealing to behavioral evidence and computational models from perceptual psychology to develop their theories of perceptual content.

[45]See Stone (2011, p 105) for an account of how visual area V5 computes motion.

[46]Interestingly, Dokic (2015), section 9, argues that our capacity for position constancy suggests that we should *reject* the existence of an egocentric coordinate space *pace* Wu.

[47]Many philosophical proponents of the VMS suggest that it should be used to explain human performance on *visuo-motor tasks* (e.g., saccading, tracking, reaching, grasping, catching, rotating, etc.). For example, see Wu (2014, p 399), Dokic (2015, p 454), Burge (2003, p 330), Hatfield (2009, p 205), Peacocke (1992, p 94). However, none of these five philosophers present an argument for a VMS based on its alleged ability to explain visuo-motor behavior. Nanay (2013, pp 39-42) offers a more detailed analysis of visuo-motor behavior, and notably he does not espouse a VMS in his theory of action-guiding representations (despite positing representations of an object's size, shape, location, etc.).

This cross-disciplinary approach makes them vulnerable to attacks on the basis of new behavioral evidence. In this section, I will discuss two sets of experimental results that I believe undermine the existence of a VMS. These experiments test human performance on *distance report tasks* and *visual crowding tasks*.

The first set of experimental results suggest that, in many environments, humans perceptually represent distance relations asymmetrically. That is, the perceived distance from point A to point B is often systematically unequal to the perceived distance from point B to point A. If this is right, the symmetry axiom of a metric space would be violated in spatial perception. I am not familiar with any VMS proponents that have addressed this challenge to their view. Two experiments revealing asymmetric distance estimations are presented in Codol (1990). In the first experiment, seventy-two participants were each placed in a rectangular room with twelve letter-labeled wooden disks placed on the floor at variable distances from one other (p 395). Each participant was asked to stand on one disk and answer (in centimeters) questions of the form: (i) "How far are you from person P?," and (ii) "How far is person P from you?" (p 394). The results show that, on average, humans estimate the distance in (i) to be significantly greater – more than 14% greater – than the distance in (ii). The second experiment confirmed this asymmetry in perceived distance. In this experiment, Codol gave participants a map of a public room containing letter-labeled points. One point was labeled "ME," and the participants were told the other points represented people. Participants were asked to answer questions (i) and (ii) again (using a map legend showing that one meter corresponds to a 3-cm line). Their answers confirmed the original results. Codol concludes that "individuals tend to consider others closer to themselves than they consider themselves to others" (1989, p 15). Could a VMS proponent account for these results? It is certainly possible, but to do so she would have to claim that these distance reports are not based on the VMS.[48] The simplest explanation, to my mind, is that these distance estimates are based on perceptual attributions that do not constitute a metric space.[49]

The other set of experimental results comes from the literature on *visual crowding*. Visual crowding can be characterized as any phenomenon in which humans are unable to individuate, categorize, or order objects in peripheral vision when they are flanked by other objects. Vision scientists usually distinguish visual crowding phenomena from the limitations associated with reduced peripheral acuity:

> Visual crowding refers to the phenomenon in which a target may be easily recognizable when viewed in isolation but becomes difficult to identify when flanked by other items. The ease of recognizing the isolated target indicates that crowding is not simply a by-product of reduced visual acuity in the periphery. Instead, it seems that the visual system applies some as-yet-unspecified lossy transformation – perhaps some form of "feature integration," pooling, or averaging – to the stimulus, resulting in the subjective experience of mixed-up, jumbled visual features. (Balas et al. 2009, p 1, citations removed)

One diagram from Rosenholtz (2016, fig 4) provides a nice example visual crowding. In the first and third rows, observers fixate on the cross and they are able to categorize the 'V' on the left as a V. By contrast, in the second row, "an observer might see these crowded letters in the wrong order ... they might not see a V at all or might see strange letter-like shapes made up of a mixture of parts from several letters" (pp 443-444). Similarly,

[48] Peacocke (1992) may be providing a response to this challenge when he distinguishes VMS perceptual distance content and non-VMS perceptual distance content. He admits that one may perceive "that x is above y, that y is above z, and that z is above x" (79), and I presume that he thinks similar paradoxical perceptions occur when we represent quantitative distance. To account for this fact, he claims that these judgments derive *not* from his VMS ("scenario content") but from a different type of perceptual content that he calls "protopropositional content." This is certainly a possible way out of the predicament, but this raises the question of why we would need to posit is a VMS at all.

[49] Similar results have been established for spatial memory (rather than perception). See Newcombe *et al.* (1999) for an overview. These results potentially constitute evidence against the existence of a *cognitive map*, but they do not directly say anything about the contents of perception.

Rosenholtz (2015) notes that an "observer might see these crowded letters in the wrong order, perhaps confusing the word ['BOARD'] with 'BORAD' " (p 11).

Many vision scientists have used these behavioral results to develop 'bag-of-features' models of visual perception (see Pelli *et al.* 2004 for a review). According to these models, our visual field is split up into "pooling regions" that can be used to index low-level features such as contours and textures *before* we individuate, categorize and order visible objects. These models provide clear explanations for many of the perceptual failures that humans exhibit in visual crowding tasks. If these models are correct, we likely do not perceive *most* of the spatial relations that hold between objects even when we have perceptually individuated these objects. If the visual perception systematically fails to preserve even the *order* of visible objects, then it is unlikely that the much stricter axioms of a metric space are satisfied by the contents of visual perception. I believe visual crowding phenomena are much more easily incorporated into a theory of perceptual content that forgoes a VMS in favor of piecemeal spatial relations.

The above experimental results associated with distance report tasks and visual crowding tasks constitute anomalies for VMS theories of spatial perception. These results are not decisive refutations of the VMS view, but I believe they should be taken seriously by any philosopher who thinks visual perception has a unified geometry. In the second part of this essay, I will offer an overview of what I take to be the current dominant paradigm of spatial perception in vision science. I see this paradigm as constituting a non-geometric alternative to the VMS framework for understanding spatial perception that currently dominates philosophical theorizing.

1.3 Spatial Representation according to Vision Science

In the first part, I attempted to show how the major philosophical arguments for a VMS are flawed. These arguments – from object shape categorization (Peacocke, Burge), scene categorization (Peacocke), demonstrative reference (Strawson, Evans), navigation (Burge), size estimation (Hatfield), and motion and position constancy (Wu) – all fail to support the belief in a VMS.[50] In this part, I would like to address the argument from psychological authority for a VMS. In particular, most philosophical proponents of the VMS assume that the psychological research of David Marr provides scientific legitimacy to the VMS. Evans, Peacocke, and Hatfield all explicitly draw on Marrs 2.5-D sketch for inspiration in developing their VMS variants. In section 3.1, I will show why Marr's views and arguments provide no support for the VMS. In section 3.2, I will show that any attempt to revise this argument from authority by replacing Marr with the "best current vision science" provides no support for positing a VMS. In section 3.3, I will conclude by discussing our prospects for understanding spatial *memory* and the nature of its geometric content.

1.3.1 Marr's visual perception: demarcation and primary function

Many philosophers see David Marr as a scientific authority on visual perception. However, in the context of arguing for a VMS, Marr is not an ideal inspirational source for three reasons. First, it is questionable whether Marr even posits a VMS. Second, Marr's reasons for positing the 2.5-D sketch are not conceptually valid. Third, the current best empirical theories of the representational content underlying the visual tasks that Marr posits his 2.5-D sketch to explain – namely, shape and scene categorization – do not appeal to a metric space. I will develop these three points below, but it is necessary to first outline which parts of Marr's

[50] Additionally, I am unaware of any more recent experimental results that purport to show that humans make use of a VMS in completing any of these six visual tasks.

approach to visual perception were innovative in the years around 1980 and remain dominant in vision science today: (i) his isolation of the computational level, (ii) his demarcation of perception from sensation, and (iii) his view of the primary function of the visual system.

Marr and Poggio (1977) introduced the notion of a "computational level" to vision science as a way of highlighting the limitations of a 'neurons-up' approach to vision science. The computational level of analysis uses goal-oriented and representational terms to describe visual processes. According to Marr, this terminology is most often used by the plain man and the computer vision engineer (1982, p 4). In contrast, brain scientists like Barlow (1972) reject this terminology and focus exclusively on "how the nervous system is built and how parts of it behave ... how the cells are connected, [and] why they respond as they do" (Marr 1982, p 4). The plain man and the computer vision engineer, however, understand visual processes primarily as an attempt to solve certain representational *problems*, such as stereopsis, color attribution, perceptual grouping, and shape attribution (1977, pp 476-481). According to Marr, the neurons-up approach of Barlow is unable to distinguish hexagonal after-images from the interpretations of the Necker cube (p 471). The latter, but not the former, is intrinsically a representational and goal-oriented process. Like other representational problems, a scientific understanding of how we perceive the Necker cube will require a computational level of analysis in addition to the neuron-based hardware level of analysis (1982, p 25). Vision science today accepts the need for a computational level of analysis. Many articles in perceptual psychology continue to frame their experiments and discussions as an analysis of specific representational problems.

Marr's second innovation was his general distinction between sensation and perception that aligns with the distinction I offered in section 1.1. In particular, Marr appeals to *constancy* transformations as a feature of perceptual states that demarcates them from purely sensory states:

> The important thing about the senses is that they are channels for perception of the real world outside... [J. J. Gibson (1966)] asked the critically important question, How does one obtain constant perceptions in everyday life on the basis of continually changing sensations? This is exactly the right question, showing that Gibson correctly regarded the problem of perception as that of recovering from sensory information "valid" properties of the external world. (1982, p 29)

To illustrate this border between sensation and perception, Marr offers a few examples of visual processes that never "obtain constant perceptions": the housefly landing on the ceiling (pp 32-34), and the human eye making a saccade in response to a sudden environmental change (p 105). Marr sees himself (and Gibson) as unlike most mid-century perceptual psychologists who "have made no serious attempts at an overall understanding of what perception is, concentrating instead on the analysis of properties and performance" (p 9). These psychologists simply focus on, e.g., color receptors, shape judgments, or stereogram-based judgments. Marr and Gibson want to give a synoptic account of perception. They both adopt a framework in which perspective-invariant constancies demarcate perceptual states from purely sensory states. This much is accepted by mainstream vision science today.[51]

Mar's third lasting contribution to vision science is his claim that the *primary function* of the human visual system is representing constancies of shape and spatial layout:

> The quintessential fact of human vision [is] that it tells about shape and space and spatial arrangement. Here lay a way to formulate its purpose – building a description of the shape and positions of things from images. Of course, that is by no means all that vision can do; it also tells about the illumination and about the reflectances of the surfaces that make the shape – their brightnesses and colors and visual textures – and about their motion. But these things seems

[51] Cf. footnote 7

secondary; they could be hung off a theory in which the main job of vision was to derive a representation of shape. (1982, p 36)

Marr says that he formed this view of the primary function of human vision on the basis of Elizabeth Warrington's experiments in clinical neurology, which show that objective object shape categorization can occur without semantic categorization (Warrington and Taylor, 1973). More specifically, Marr goes on to claim that the primary function of the human visual perception is attributing shape to various particulars: portable objects, agents, faces, scenes, and surfaces.[52] This shape-based view of the *primary* function of visual perception informs the rest of 52 Marr's book and continues to be the dominant framework in vision science today (Kanwisher and Dilks, 2014). In sum, Marr's lasting contribution to vision science is the proposition that the primary function of visual perception is to represent perspective-invariant shape properties via constancy transformations.

Other aspects of Marr's work have not lasted. In particular, Marr proposes a set of syntactic-computational *stages* for recovering objective shape representations, most of which are no longer seriously endorsed by vision scientists. Marr's stages are: the retinal image, the zero-crossing, the raw primal sketch, the full primal sketch, the 2.5-D sketch, and the 3-D model. Computational theories today usually only endorse the first two stages.[53] Given that philosophers continue to draw on Marr's 2.5-D sketch for inspiration for their VMS (Peacocke 1992, p 65; Hatfield 2009, p 5), it is important to note that the 2.5-D sketch is no

[52]For objects, see pp 35, 312, 316. For agents, see p 319. For faces, see p 311. For scenes, see pp 217, 240, 249 (and arguably pp 229-232). For object surfaces, see pp 220-221, 224, 229. Note that most vision scientists and philosophers use "shape perception" in a restricted sense to mean perceptual attribution of shape to manipulable objects. For Marr, shape perception will also include aspects of what today is often called "body perception," "face perception," "perceptual organization," "spatial layout perception," and "scene perception."

[53]Carandini *et al.* (2005), Freeman and Simoncelli (2011). Even in the Forward to the 2010 edition of Marr's *Vision*, Marr's former student Shimon Ullman admits that the currently dominant approach to object shape perception no longer posits a VMS such as Marr's 2.5-D sketch or 3-D model: "Computational vision has been dominated in the last decade by an alternative approach to [object 3D shape] recognition, based on describing the possible image appearances of an object rather than its invariant 3-D structure" (xxi). Ullman is here referring to the HMAX model of Reisenhuber and Poggio (1999) that I mentioned in footnote 40 above.

longer a staple of computational theories of visual perception.[54] Another aspect of Marr's approach to vision that has not lasted is his attempt to integrate representational theories from vision science and visual psychophysics.[55] The latter tradition accepts the existence of a VMS on phenomenological and introspective grounds.[56] It posits a coordinatized geometry for the VMS on the basis of experimental estimations of apparent size and distance. Marr briefly embraces this phenomenological methodology when he appeals to apparent distance and orientation estimationresults in Ittelson (1960) and Stevens (1979) to justify the claim

[54]The 2.5-D sketch is defined as an encoding of "properties of the visible surfaces in a viewer-centered coordinate system, such as surface orientation, distance from the viewer, and discontinuities in these quantities" (p 38).

Although my line of reasoning here does not depend on it, I think the case could be made that Marr is committed to view that the 2.5-D sketch is *not* a representation at all, but is rather a merely syntactic entity. This is because representations have to be "objective," "physical," and "permanent" (pp 29, 34, 36). Hence,the fact that the 2.5-D sketch is "viewer-centered" seems to disqualify it from being a representation. According to Marr, the housefly has "no explicit representation" (35) of its landing surface because its visual system makes only "subjective measurements" and does not describe "objective qualities" (36). These claims seem to be true of the 2.5-D sketch as well, and hence it should not be understood as a representation. Against his better judgment, Marr nevertheless claims that the 2.5-D sketch represents "objective physical reality" (p. 269). Given Marr's lack of clarity on this matter of whether *subjective but perspectivally semi-invariant properties* (such as egocentric distance, direction, or slant) can be perceptually represented, I will bracket the question of what Marr's considered view is on whether his 2.5-D sketch is representational.

[55]In vision science, Marr draws on Hubel and Wiesel (1968). In visual psychophysics, Marr draws on Ittelson (1960) and Stevens (1979). For the sake of clarity, I understand "vision science" to refer to any fixation-controlled experimental results (behavioral or neural), as well as the theoretical claims made by these experimentalists in discussion sections, review articles, meta-analyses, and textbooks. More specifically, I understand *vision-scientific behavioral results* to include data from fixation-controlled experimental tasks such as forced choice, categorization, estimation, serial reidentification, serial discrimination, instantiation identification, navigation, reaching, etc. (These results may or may not depend on neural atypicalities as a TMS or an agnosia.) I have in mind experimentalists such as Mary Potter, Ruth Rosenholtz, and Aude Oliva. I understand *vision-scientific neural results* to include fMRI, EEG, and ECoG data about adaptation, differential regional blood oxygen level-dependent (BOLD) activation across identical stimuli, and intra-regional BOLD activation across differing stimuli. I have in mind experimentalists such as Nancy Kanwisher, Russell Epstein, and Daniel Dilks.

Visual psychophysics, however, is a collection of experimental results in which participants estimate the *apparent* (rather than *objective*) metric size and distance features. On this basis, visual psychophysics articles infer that their postulated VMS exhibits various coordinatized geometries (Henderson 19999, Indow 2004, Wagner 2006, Koenderink and van Doorn 2008). This tradition and these psychologists are the ones that Hatfield (2003, 2009) draws on as his scientific authorities in his argument for a contracted Euclidean geometry for visual space.

[56]Additionally, many theorists in the visual psychophysics tradition accept a sense-data framework for perceptual content. For instance, Ittelson (1960, p 133) claiming that "apparent distance is a subjective datum which cannot be observed or measured directly by the experimenter." Wagner (2006, ch 1) makes many similar claims. These theorists presuppose a veil of perception in which we attribute locations and properties to phenomenal entities. As I mentioned in section 1 while discussing Hatfield, I strongly suspect that these phenomenological reports are based on higher cognitive processes (e.g., conversational maxims). To suppose otherwise requires ascribing increasingly esoteric and context-dependent geometric structure to the content of perception (cf. Wagner 2006).

that "there is at least one internal representation [i.e., the 2.5-D sketch] of the depth [i.e., egocentric distance], surface orientation, or both associated with each surface point in a scene" (1982, pp 275, 283).[57] Marr's argument for a VMS on the basis of visual psychophysics would not be accepted in vision science today because appearance estimation results are a poor substitute for the adaptation results and fixation-controlled behavioral results that are currently used to determine perceptual contents.[58] But even if one were to follow Marr in embracing the validity of this phenomenological methodology, it is notable that Marr himself admits that the psychophysical evidence for the existence of his 2.5-D sketch is weak:

> Unfortunately, I cannot provide much more than a framework within which to ask questions. ...There has not yet been any determined psychophysical assault on the 2.5-D sketch, so we know very little about it or even whether it in fact exists in the sense suggested by our approach to vision. (1982, p 279)[59]

Despite Marr's cautionary claims about the existence of his 2.5-D sketch, philosophers and psychologists have often assumed that its existence is empirically well-supported. This is simply not the case. Hence, a legitimate argument from authority for a VMS should not be appealing to Marr's writings on the 2.5-D sketch.

[57]Marr offers some other arguments for a VMS, but they are quite opaque. In one argument, Marr appeals to Julesz's results that there is a 2°-of-disparity upper limit on random-dot stereoscopic fusion. He claims that this result indicates that we are likely perceiving the egocentric distance of all the surfaces in the perceived scene (Marr 1982, pp 279-282). I fail to see how this argument is supposed to work. This upper limit be a direct consequence of the physiology of matching the two retinal images in the striate cortex. But more importantly, I do not see how an upper limit on fusion *could* support a claim about the perception of egocentric distance. (One might argue that fusion itself requires a representation of egocentric distance, but this is not what Marr is claiming.) In another argument, Marr appeals to occlusion perception to justify the claim that we perceive egocentric distance. But we may simply be perceiving a *non-metric* depth relations ('in front of' and 'behind') between surfaces (cf. Bennett 2016, pp 28ff.).

[58]This is simply an observation about the kinds of evidence that vision scientists currently use to infer perceptual content. But I also believe there are good *a priori* reasons for being wary of such psychophysical evidence: (i) these experiments are not fixation-controlled and thus are unable to distinguish perceptual content from cognitive content, and (ii) the metric estimates are usually for a single surface/object and thus can reveal very little about any alleged global representation of the environmental geometry.

[59]For example, Marr claims that "we do very poorly" on judging which of two surfaces is farther away when they "lie in different parts of the visual field" (p 282). This fact "casts doubt ... on the idea that depth is ... stored accurately over a range of values" (1982, p 282). The estimation results of Norman *et al.* (1995) gives rise to a parallel doubt for surface orientation.

As we have seen in this section, Marrs lasting contribution to vision science was the idea that the primary function of the human visual system is to attribute perspective-invariant shapes to portable objects, agents, faces, scenes, and surfaces via constancy transformations. Marr's other contributions, such as his posited stages of visual processing, have not lasted. (And, at least in the case of the 2.5-D sketch, we have seen that this was for good reason.) In the next section, I will elaborate the *current* vision-scientific understanding of the spatial contents of perception. As we will see, the current explanations of our most well-understood everyday visuo-spatial tasks make no appeal to a VMS.

1.3.2 Current Views on Shape and Scene Perception

For Marr, the primary function of visual perception was shape attribution to portable objects, agents, faces, scenes, and surfaces. In many ways, this is still the dominant paradigm of visual perception. This view of visual perception has been repeatedly confirmed since Kanwisher and colleagues introduced functional neuroimaging techniques into vision science in the 1990s. They have isolated neural correlates and pathways in the ventral stream for perceptual attribution of each of the following shape types that Marr discusses: object/surface shape, face shape, body shape, and scene shape. Kanwisher and Dilks (2014) summarize the "key results from the last 15 years of neuroimaging research on humans":

> The central finding from this now-substantial body of work is that the VVP [ventral visual pathway] is not homogeneous but is instead a highly differentiated structure containing a set of regions each with its own distinct functional profile. These regions include the fusiform face area (FFA), which responds selectively to faces, the parahippocampal place area (PPA), which responds selectively to [scenes], the extrastriate body area (EBA), which responds selectively to bodies, the lateral occipital complex (LOC), which responds to object shape largely

> independent of object category, and the visual word form area (VWFA), which responds selectively to both visually presented words and consonant strings. Each of these regions is present in approximately the same location in virtually every healthy subject. These regions and their cohorts (e.g., the occipital face area or OFA) constitute the fundamental machinery of high-level visual recognition in humans. (p 733, citations removed)

These neural correlates and pathways respond selectively to their associated shape types within a few hundred milliseconds.[60] Moreover, we know that these cortical areas represent distal environmental constancies because they exhibit adaptation effects. By recording these adaptation effects, Kanwisher and colleagues infer which properties are being represented in these cortical areas.[61] I will focus on the question of which spatial contents mediate our visual achievement of *object shape constancy* and *scene shape constancy* given that (i) these are the two most commonly investigated and well-understood representational processes from the five listed above, and (ii) they are the two representational processes of the five that our philosophers appealed to in arguing for a VMS.[62]

Scene Categorization

Since 2001, Aude Oliva and her colleagues have been the leading figures in vision science attempting to answer the question: what are the spatial contents involved in perceiving an environmental space (i.e., a scene)? They claim that a decade of fMRI adaptation results have suggested that "an environmental space can be represented by two separable and complementary descriptors (Oliva and Torralba, 2001): its spatial boundary (i.e., the shape and

[60] See Bastin et al. (2013), Greene and Oliva (2009a) for the temporal profiles of PPA and OPA.

[61] In addition to fMRI adaptation, visual neuroscientists infer representational content based on the neural and behavioral results described in footnote 54.

[62] The totality of these philosophical arguments were based on: object shape categorization (Peacocke, Burge), scene categorization (Peacocke), motion and position constancy (Wu), demonstrative reference (Strawson, Evans), navigation (Burge), and size estimation (Hatfield). I will address the last four at the end of section 2.2.

size of the scene's space) and its content (textures, surfaces, materials and objects)" (Park et al. 2011, p 1333). The 'spatial boundary descriptor' attributes global features to the environmental space, whereas the 'content descriptor' attributes local features to various items in the environment. These two perceptual representations arise in two distinct visual pathways that end in the "PPA and LOC," respectively (p 1339). I will discuss the representational contents of the latter pathway below. In the former pathway, these "scene-selective" visual areas (e.g., OPA, PPA, RSC) exhibit adaptation effects when we add or remove portable objects to the scene over multiple fixations, but these areas do *not* exhibit such adaptation effects when we alter the permanent boundaries and landmarks of the scene (Park and Chun 2009; Park *et al.* 2011).

What, then, is the content of Oliva's proposed "spatial boundary" percept associated with the PPA? According to Oliva's model, this content is given by a *global feature space*.[63] For example, the *openness* and *depth* of the visible environment – defined in terms of perspective-invariant large-scale texture patterns within the retinal image (see figure 3 of Torralba and Oliva (2003), p 395) – are perceived attributes of this space.[64] These features may then subserve various visual tasks such as categorizing scenes as a "beach," "street," or "forest."

Although the exact number of global features that we represent in this pathway is not yet known, the fMRI adaptation results are consistent with the five global features postulated in Oliva's model (the *Spatial Envelope Model*).[65] Significantly, neither Oliva's model nor its

[63] Park *et al.* (2015), Oliva *et al.* (2011), Greene and Oliva (2006, 2009a,b), Oliva and Torralba (2001).

[64] For instance, Greene and Oliva (2009a), Torralba and Oliva (2003), and Oliva and Torralba (2001) use various gabor filters (a type of fourier transformation) on an image to classify the visible environment along feature dimensions such as *naturalness, openness, roughness, expansion,* and *ruggedness.* The neural location of these spatial boundary attributions may be in visual areas adjacent to the PPA. For instance, the RSC seems to exhibit greater perspective-invariance than the PPA along some dimensions: "We did not find any attenuation for panoramic repeats in the PPA, showing viewpoint-specificity. In contrast, RSC showed significant attenuation for the panoramic condition, showing viewpoint-integration." (Park and Chun 2009, p. 1747)

[65] See Oliva (2014), pp 727ff., for references to some of these adaptation experiments. For additional studies on the constancies exhibited in the visual pathway to PPA, see Dilks *et al.* (2013), Kamps *et al.* (2016), Dilks *et al.* (2011), Greene and Oliva (2009b), Park and Chun (2009), Ward *et al.* (2010), Morgan

competitors posit a VMS: the distance relations between visible objects are never represented as such.[66] Many vision science articles are explicit on this point, such as this illustrative passage from Oliva in which she contrasts her "holistic-based approach" to scene perception with Marr's "part-based approach":

> In the attempt to explain how the brain represents a scene, the part-based approach (Marr, 1982; Biederman, 1987) depicts access to scene meaning as the last step within a hierarchical organization of modules of visual processing with increasing complexity (edges, surfaces, objects, scene). The "geon" theory put forth by Irving Biederman (1987, 1995) suggests that fast scene understanding could be achieved via a representation of the [geometric] arrangement of simple volumetric forms from which the identity of the individual objects and scenes can be inferred. Alternatively, a holistic-based approach (spatial envelope theory; see Oliva and Torralba, 2001) constructs a meaningful representation of scene gist directly from the low-level features pool, without binding contours to form surfaces, and surfaces to form objects. (Oliva 2005, p 256)[67]

et al. (2011), Epstein *et al.* (2007). These studies suggest that spatial layout may be represented in terms of spatial frequency textures as well as mean depth, center of mass, and mean element size. Although I cannot explore these models here, it is notable that none of these features suggest the existence of a VMS.

[66]The competing models of scene representation in vision science are: (i) the *Texture Tiling Model* of Rosenholtz (2009, 2012, 2016), (ii) the *Parametric Texture Model* of Gatys *et al.* (2015) and Portilla and Simoncelli (2000), (iii) the *Scene Schema Model* of Renninger and Malik (2004), Kim and Biederman (2012), and potentially Dillon *et al.* (2017), and (iv) the *Regional Semantics Model* of Vogel and Schiele (2007). All four of these models are very similar to Oliva's model in that they appeal to texture patterns (i.e., spatial frequencies) on the retinal image – rather than a metric space – to explain human performance on visual tasks such as categorizing and reidentifying scenes in rapid serial visual presentations. Only the *Scene Schema Model* posits anything like a geometric structure in representation content, but even here it is a topological relational structure in which the relations are analogous to locative prepositions such as "under" and "on." These more abstract, topological relational structures were first posited by Biederman as "scene schemas" (1982, p 145) and "structural descriptions" (1987, p 128). I find it somewhat plausible that these structures exist, but there are few neural or behavioral results that support their existence in single-fixation perception.

[67]For similar passages, see Mullin and Steeves (2011), p 4174; Mullally and Maguire (2011), p 7441; Kanwisher and Dilks (2014), pp 736-737; Rosenholtz (2016), p 449.

As we saw in section 2.1, Marr claims that the visual system first encodes local surface properties of egocentric distance and orientation in a 2.5-D sketch and only later do we end up perceiving the spatial layout or semantic properties of a scene. For Oliva and most present-day vision scientists, Marr's approach is not only an incorrect account of the computational stages preceding scene perception, but it is an incorrect account of the *representational contents* of scene perception. Marr postulated that the shape we perceptually attribute to a scene will necessarily be embedded in a coordinate system in which each visible point of a surface receives its own local coordinate. Oliva rejects this "part-based approach" in favor of an approach in which spatial layout contents are analyzed in terms of global spatial frequencies, not coordinate systems. In other words, the dominant and empirically-supported approach to scene perception today outright rejects the postulation of a 2.5-D sketch or a VMS. The current models claim that the *naturalness, openness, roughness, expansion, ruggedness, etc.* of the visible environment are each attributed to the global space without first locating multiple objects within a metric space.[68]

In sum, the best current models of scene perception do not posit a VMS, and many articles explicitly reject the VMS as an artifact of a now-defunct tradition in vision science. These current models are supported by a substantial body of neural and behavioral evidence, almost all of which was not available to Marr in the 1980s. Today, a scientifically-informed account of scene perception all but mandates that we reject a VMS in favor of a global feature space representation. Such global features are stable over changes in perspective and many studies show that various visual areas in the pathway to PPA adapt to these features.

[68] As might be guessed, Oliva considers her model to be Gibsonian in spirit: "By grounding our search in the principles of environmental affordance (Gibson, 1979; Rosch, 1978), we have been able to find a collection of properties that are necessary and sufficient to capture the essence of many landscape image categories" (Greene and Oliva 2006, p 6).

Object Shape Categorization

Does the VMS-concept fare any better in the study of object shape perception? That is, do any of the "object-selective regions" in the LOC contain a VMS? Neural evidence suggests otherwise: these regions exhibit adaptation across changes in the object's occlusion, size, position, handedness, surface curvature, and internal angles.[69] If object shape contents in the LOC included a metric space (with metric relations been a set of vertices, for instance), these adaptation effects would not occur since the VMS itself would vary with each of these changes in the object's features.[70]

A VMS proponent may wonder whether these abstract and non-metric shape contents in the LOC might be constructed from a VMS representation that arises in a pre-LOC visual area such as V4. Indeed, vision scientists have shown that when "identical stimuli differing in size ... were presented, V4 showed a recovery from adaptation" (Konen and Kastner 2008, p 227). On its own, this could be used to suggest that V4 represents a metrically-determined shape of an object. However, this size-dependency of V4 taken in conjunction with the fact that V4 *also* recovers from adaptation when the object's *orientation* changes suggests that "the previously observed adaptation effects [in V4] were due to adaptation to low-level features" of the retinal image (p 227). Hence, V4 may not exhibit any perceptual constancies, "whereas higher-order lateral occipital complex (LOC) responds selectively to objects inde-

[69] For changes in occlusion, see Kourtzi and Kanwisher (2001). For changes in size and position, see Grill-Spector *et al.* (1999). For changes in handedness, see Dilks *et al.* (2011). For changes in surface curvature and internal angles, see Kourtzi *et al.* (2003) as well as the studies of the LOC-homologue in rhesus monkeys in Kayaert *et al.* (2003), Kayaert *et al.* (2005). (The dishabituation results in Kayaert and Wagemans (2010) provide behavioral corroboration that affine rather than metric features are the primary contents of shape perception.) Somewhat relatedly, the results of Kim and Biederman (2012) suggest that we perceive spatial relations between *multiple* objects (e.g., attachment, alignment, occlusion) only if these relations are non-metric.

[70] The parts and structure of these non-metric shape representations are still not well understood. There are various programmatic frameworks, however, such as Biederman's geon theory (Biederman 1987; Lescroart and Biederman 2013). One might wonder whether we know that the LOC represents object *shape* at all rather than, say, an object's even more abstract semantic properties. But the latter view has been disconfirmed by the adaptation results mentioned in footnote 68 as well as by a study that directly tested this view: "fMRI adaptation is not found in LOC across objects that are similar in meaning but differ in shape" (Kanwisher and Dilks 2014, p 736).

pendent of image transformations, suggesting a more abstract visual representation that is necessary for perceptual object constancy" (Konen *et al.* 2011, p 49).[71]

Perceptual Content Other Than Shape-Based Content

In sum, the evidence from vision science over the past twenty years suggests that there is no VMS in the representational content underlying object shape and scene shape perception. Hence, it would seem that any philosophical attempt to update Peacocke's or Burge's appeals to object shape or scene shape perception to argue for a VMS would be fighting an uphill battle against the current vision-scientific consensus. Of course, as we saw in Part 1, there are other kinds of perceptual content on the basis of which one might argue for a VMS. Wayne Wu appeals to motion perception, Peter Strawson and Gareth Evans appeal to perceptual reference, Burge appeals to perceptual content used in navigation, and Hatfield appeals to perceptual content used in size estimation judgments. As we saw in Part 1, the particular arguments for a VMS that these authors provide are flawed. Does current vision science provide any experimental results about these topics that could be used to update and strengthen their arguments? If so, I am unaware of such results. Although there have been sustained vision-scientific investigations of perceptual reference (in the form of object tracking systems), object motion, and navigation, but little is known about the spatial contents in these perceptual systems.That is, vision scientists currently lack experimentally-grounded theories about which spatial features these perceptual processes actually *represent*.[72] Hence, it would seem that appealing to vision-scientific authority regarding *any* of these perceptual contents, shape-based or not, will *at best* be unhelpful to a VMS proponent.

[71]For a similar claim that visual areas before LOC do not represent distal properties (i.e., invariances, constancies), see Kanwisher and Dilks (2014, p 737). For fMRI adaptation results that suggest V4 may exhibit constancy in representing the metric distance between two surfaces, see Neri *et al.* (2004), p 1881.

[72]For the little we do know about the spatial contents in these systems, see the following articles. Object motion: Ashida *et al.* (2007), Warren and Rushton (2007), Fajen and Matthis (2013). Perceptual reference: Xu (2009), Green (2018). Navigation: Raudies and Hasselmo (2011).

1.4 Conclusion and Open Questions

Many philosophers in the 20th and 21st centuries have claimed that humans represent a metric space in visual perception. I have critiqued their central arguments for this claim in section 2, and I have contrasted this claim with the most experimentally-grounded theories of perceptual content in vision science today in section 3. Based on these considerations, I see every reason to be skeptical about the existence of a VMS. At the same time, humans constantly rely on geometrically-organized cognitive representations of both physical domains (e.g., cognitive maps) and ideal domains (e.g., Euclidean figures). If they do not arise in perception, which cognitive systems contain such geometric representations, what is their specific content, and on what basis are they constructed? Answers to these questions lack the kind of robust empirical support that investigations of perceptual content could offer. Nevertheless, these questions are vital to any basic understanding of animal and human representation. For this reason, developmental psychologists like Elizabeth Spelke and philosophers like Marcus Giaquinto have proposed representational theories to account for various kinds of geometric cognition. In the next two chapters, I draw from and criticize their theories in developing my own understanding of cognitive maps and elementary geometric beliefs.

1.5 Bibliography

Arcaro, M. J., McMains, S. A., Singer, B. D., and Kastner, S. (2009). "Retinotopic organization of human ventral visual cortex." *Journal of Neuroscience*, 29 (34), pp 10638-10652.

Ashida, H., Lingnau, A., Wall, M. B., and Smith, A. T. (2007). "fMRI Adaptation Reveals Separate Mechanisms for First-Order and Second-Order Motion." *Journal of Neurophysiology* 97, pp 1319-1325.

Balas, B. J., Nakano, L., Rosenholtz, R. (2009). "A summary-statistic representation in peripheral vision explains visual crowding." *Journal of Vision* 9(12):13, pp 1-18.

Baldwin, T. (1992). "The projective theory of sensory content." In: T. Crane (Ed.), *The Contents of Experience*. Cambridge: Cambridge University Press, pp 177-195.

Bar, M., and Aminoff, E. M. (2003). "Cortical analysis of visual context." *Neuron* 38, pp 347-358.

Barlow, H. B. (1972). "Single units and sensation: a neuron doctrine for perceptual psychology?" *Perception* 1, pp 371-394.

Bayne, T. (2016). "Gist!" *Proceedings of the Aristotelian Society* 2(116), pp 1-20.

Bell, A. H., Pasternak, T., and Ungerleider, L. G. (2014). "Ventral and Dorsal Cortical Processing Streams." In Werner and Chalupa (2014), pp 227-242.

Bennett, D. J. (2011). "Seeing Shape: Shape Appearances and Shape Constancy." *British Journal for the Philosophy of Science* 0, pp 1-32.

Biederman, I., Mezzanotte, R. J., Rabinowitz, J. C. (1982). "Scene Perception: Detecting and Judging Objects Undergoing Relational Violations." *Cognitive Psychology* 14, pp 143-177.

Biederman, I. (1987). "Recognition-by-components: A theory of human image interpretation." *Psychological Review* 94, pp 115-148.

Block, N. (2012). "The Grain of Vision and the Grain of Attention." *Thought* 1, pp 170-184.

Block, N. (2013). "Seeing and Windows of Integration." *Thought* 1, pp 1-11.

Block, N. (2014a). "Seeing-as in the Light of Vision Science." *Philosophy and Phenomenological Research* 89(3), pp 560-572.

Block, N. (2014b). "Rich Conscious Perception Outside Focal Attention." *Trends in Cognitive Sciences* 18(9), pp 445-447.

Boghossian, P. A., and Velleman, D. J. (1989). "Color as a secondary quality." *Mind* 98, pp 81-103.

Briscoe, R. E. (2009). "Egocentric Spatial Representation in Action and Perception." *Philosophy and Phenomenological Research* 79(2), pp 423-460.

Briscoe, R. E. (2011). "Mental Imagery and the Varieties of Amodal Perception." *Pacific Philosophical Quarterly* 92, pp 153-173.

Burge, T. (2010). *Origins of Objectivity.* Oxford: Clarendon Press.

Campbell, J. (1994). *Past, Space, and Self.* Cambridge: MIT Press.

Carandini, M., Demb, J. B., et al. (2005). "Do We Know What the Early Visual System Does?" *The Journal of Neuroscience* 25(46), pp 10577-10597.

Carruthers, P. (2016). "Higher-Order Theories of Consciousness." *Stanford Encyclopedia of Philosophy.*

Cavanagh, P. (2001). "Seeing the Forest But Not the Trees." *Nature Neuroscience* 4(7), pp 673-74.

Cavanagh, P., He, S., and Intriligator, J. (1999). "Attentional Resolution: The Grain and Locus of Visual Awareness." In C. Taddei-Ferretti and C. Musio (Eds.), *Neuronal Basis and Psychological Aspects of Consciousness.* Singapore: World Scientific, pp 41-52.

Chalmers, D. (2006). "Perception and the Fall from Eden." *The Character of Consciousness.* Oxford: Oxford University Press, 2011. Originally published in: Gendler, T. and Hawthorne, J. (eds.), *Perceptual Experience.* Oxford: Oxford University Press, 2006.

Clark, A. (2000). *A Theory of Sentience.* Oxford: Oxford University Press.

Codol, J., Jarymowicz, M., Kaminska-Feldman, M., and Szuster-Zbrojewicz, A. (1989). "Asymmetry in the estimation of interpersonal distance and identity affirmation." *European Journal of Social Psychology* 19, pp 11-22.

Codol, J. (1990). "Studies on self-centered assimilation processes." In: J.-P. Caverni, J.-M. Fabre, and M. Gonzalez (Eds.), *Cognitive Biases. Amsterdam North-Holland Press*, pp 387-400.

Cohen, J. (2015). "Perceptual Constancy." In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception.* Oxford: Oxford University Press.

Derrington, A. M., Krauskopf, J., Lennie, P. (1984). "Chromatic mechanisms in lateral geniculate nucleus of macaque." *Journal of Physiology* 357(1), pp 241-265.

De Valois, R. L. and De Valois, K. K. (1988). *Spatial Vision.* Oxford: Oxford University Press.

Dilks, D., Julian, J. B., Kubilius, J., Spelke, E. S., and Kanwisher, N. (2011). "Mirror-Image Sensitivity and Invariance in Object and Scene Processing Pathways." *The Journal of Neuroscience* 31(31), pp 11305-11312

Dilks, D., Julian, J. B., Paunov, A. M., and Kanwisher, N. (2013). "The occipital place area (OPA) is causally and selectively involved in scene perception." *Journal of Neuroscience* 33(4), pp 1331-1336.

Dillon, M., Perischetti, A. S., Spelke, E. S., and Dilks, D. D. (2017). "Places in the Brain: Bridging Layout and Object Geometry in Scene-Selective Cortex." *Cerebral Cortex* 28(7), pp 2365-2374

Dokic, J. (2015). "Perception and Space."In: Matthen, M. (ed.) *The Oxford Handbook of Philosophy of Perception.* Oxford: Oxford University Press, pp 441-458.

Domini, F., and Braunstein, M. L. (1998). "Recovery of 3-D Structure from motion is neither euclidean nor affine." *Journal of Experimental Psychology: Human Perception and Performance* 24(4), pp 1273-1295.

Dyer, F. C. (1998). "Spatial Cognition: Lessons from Central Place Foraging Insects." In R. Balda, I. Pepperberg, and A. Kamil (Eds.), *Animal Cognition in Nature.* San Diego: Academic Press, 1998.

Epstein, R. A. and Kanwisher, N. (1998). "A cortical representation of the local visual environment." *Nature* 392, pp 598-601.

Epstein, R. A. (2011). "Cognitive Neuroscience: Scene Layout from Vision and Touch." *Current Biology* 21(11), pp 437-438.

Epstein, R. A. (2014). "Neural Systems for Visual Scene Recognition." In Kveraga and Bar (2014), pp 105-134.

Epstein, R. A., Patai, E. P., Julian, J. B., and Spiers, H. J. (2017). "The cognitive map in humans: spatial navigation and beyond." *Nature Neuroscience* 20(11), pp 1504-1513.

Evans, G., and Altham, J. E. J. (1973). "The Causal Theory of Names." *Proceedings of the Aristotelian Society*, Supplementary Volumes 47, pp 187-225.

Evans, G. (1982). *The Varieties of Reference.* Oxford: Oxford University Press.

Evans, G. (1985). "Molyneuxs Question." In A. Phillips (Ed.), *Gareth Evans: Collected Papers.* Oxford: Clarendon Press.

Fajen, B. R., and Matthis, J. S. (2013). "Visual and Non-Visual Contributions to the Perception of Object Motion during Self-Motion." *PLoS ONE* 8(2), e55446.

Ferrara, K. and Park, S. (2016). "Neural representation of scene boundaries." *Neuropsychologia* 89, pp 180-190.

Fleming, R. and Anderson, B. (2004). "The perceptual organization of depth." In L. Chalupa and J. Werner (Eds.), *The Visual Neurosciences.* Cambridge: MIT Press.

Fodor, J. (1983). *Modularity of Mind.* Cambridge: MIT Press.

Foley, J. M. (1972). "The sizedistance relation and intrinsic geometry of visual space: Implications for processing." *Vision Research* 12, pp 323-332.

French, R. (1987). "The geometry of visual space." *Noûs* 21, pp 115-33

Freeman, J., and Simoncelli, E. P. (2011). "Metamers of the ventral stream." *Nature Neuroscience* 14(9), pp 1195-1201.

Gava, L., Valenza, E., and Turati, C. (2009). "Newborns' Perception of Left-Right Spatial Relations." *Child Development* 80(6), pp 1797-1810.

Gibson, J. J. (1950). *The Perception of the Visual World.* Riverside Press.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1979).*The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin

Gillam, B. (1995). "The Perception of Spatial Layout from Static Optical Information." In Epstein and Rogers (Eds.), *The Perception of Space and Motion.* Academic Press, pp 23-67.

Gillam, B. (2011). "The Influence of monocular regions on the binocular perception of spatial layout." In Harris and Jenkin (2011), pp 46-69.

Goodale, M. A., Meenan, J. P., *et al.* (1994). "Separate neural pathways for the visual analysis of object shape in perception and prehension." *Current Biology* 4, pp 604-610.

Goodale, M. A., and Milner, A. D. (1992). "Separate visual pathways for perception and action." *Trends in Neuroscience* 15, pp 20-25.

Green, E. J. (2015). "A Layered View of Shape Perception." *British Journal for the Philosophy of Science* 0, pp 1-33.

Green, E. J. (2016). "Representationalism and Perceptual Organization." *Philosophical Topics* 44(2), pp 121-148.

Green, E. J. (2017). “Attentive Visual Reference.” *Mind & Language* 32(1), pp 3-38.

Green, E. J. (2018). “What Do Object Files Pick Out?” *Philosophy of Science* 85, pp 177-200.

Greene, M. R., and Oliva, A. (2006). “Natural Scene Categorization from Conjunctions of Ecological Global Properties.” *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Vancouver, July 2006.

Greene, M. R., and Oliva, A. (2009a). “The briefest of glances: The time course of natural scene understanding.” *Psychological Science* 20(4), pp 464-472 .

Greene, M. R., and Oliva, A. (2009b). “Recognition of natural scenes from global properties: Seeing the forest without representing the trees.” *Cognitive Psychology* 58(2), pp 137-176.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). “Differential Processing of Objects under Various Viewing Conditions in the Human Lateral Occipital Complex.” *Neuron* 24, pp 187-203.

Grünbaum, A. (1962). “Geometry, chronometry, and empiricism.” In H. Feigl and G. Maxwell (Eds.), *Minnesota Studies in the Philosophy of Science*, Minneapolis, pp 405-526.

Grünbaum, A. (1973). “Empiricism and the Geometry of Visual Space.” In A. Grünbaum (Ed.), *Philosophical Problems of Space and Time*, Boston: D. Reidel Publishing, pp 152-157.

Harris, L. R. and Jenkins, M. R. M. (2011). *Vision in 3D Environments.* Cambridge: Cambridge University Press.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., Rubin, N. (2008). “A hierarchy of temporal receptive windows in human cortex.” *Journal of Neuroscience* 28, pp 2539-2550.

Hatfield, G. (2002). “Perception as Unconscious Inference.” In D. Heyer, D. and R. Rainer Mausfeld (Eds.), *Perception and the Physical World: Psychological and Philosophical Issues in Perception.* John Wiley and Sons, pp 113-143.

Hatfield, G. (2003). “Representation and Constraints: The Inverse Problem and the Structure of Visual Space.” In Hatfield (2009a), pp 153-177.

Hatfield, G. (2009a). *Perception and Cognition.* Oxford: Oxford University Press.

Hatfield, G. (2009b). “On Perceptual Constancy.” In Hatfield (2009a), pp 178-211.

Hatfield, G. (2011). “Philosophy of Perception and the Phenomenology of Visual Space.” *Philosophic Exchange* 42(1), pp 33-66.

Hatfield, G. (2012). “Phenomenal and Cognitive Factors in Spatial Perception.” In G. Hatfield and S. Allred (Eds.), *Visual Experience: Sensation, Cognition, and Constancy.* Oxford: Oxford University Press.

Heelan, P. A. (1983). *Space Perception and the Philosophy of Science.* Berkeley: University of California Press.

Hershenson, M. (1999). *Visual Space Perception: A Primer.* Cambridge: MIT Press.

Hubel, D. H., and Wiesel, T. N. (1968). “Receptive fields and functional architecture of monkey striate cortex.” *J. Physiol.* (Lond.) 195, pp 215-243.

Indow, T. (2004). *The Global Structure of Visual Space.* World Scientific Publishing.

Intraub, H. (2014). “Visual Scene Representation.” In Kveraga and Bar (2014), pp 5-26.

Intraub, H., and Richardson, M. (1989). “Wide-angle memories of close-up scenes.” *Journal of Experimental Psychology, Learning, Memory, and Cognition* 15(2), pp 179-187.

Kant, I. (1787). *Critique of Pure Reason.* P. Guyer and A. Wood (Eds.). Cambridge: Cambridge University Press, 1999.

Kanwisher, N. (2004). “The ventral visual object pathway in humans: evidence from fMRI.” In L. Chalupa and J. Werner (Eds.), *The Visual Neurosciences.* Cambridge: MIT Press.

Kanwisher, N., and Dilks, D. D. (2014). “The Functional Organization of the Ventral Visual Pathway in Humans.” In Chalupa and Werner (2014), pp 733-746.

Kayaert, G., Biederman, I. and Vogels, R. (2003). “Shape Tuning in Macaque Inferior Temporal Cortex.” *The Journal of Neuroscience* 23, pp 3016-3027.

Kayaert, G., Biederman, I. and Vogels, R. (2005). “Representation of Regular and Irregular Shapes in Macaque Inferotemporal Cortex.” *Cerebral Cortex* 15, pp 1308-21.

Kayaert, G. and Wagemans, J. (2010). “Infants and Toddlers Show Enlarged Visual Sensitivity to Nonaccidental Compared with Metric Shape Changes.” *i-Perception* 1, pp 149-158.

Konen, C. S., and Kastner, S. (2008). “Two hierarchically organized neural systems for object information in human visual cortex.” *Nature Neuroscience* 11(2), pp 224-231.

Konen, C. S., Mruczek, R. E. B., Montoya, J. L., and Kastner, S. (2013). “Functional organization of human posterior parietal cortex: grasping- and reaching-related activations relative to topographically organized cortex.” *Journal of Neurophysiology* 109, pp 2897-2908.

Kornblith, S., Cheng, X., Ohayon, S., and Tsao, D. (2013). "A Network for Scene Processing in the Macaque Temporal Lobe." *Neuron* 79, pp 766-781.

Kravitz, D. J., Saleem, K. S., Baker, C. I., and Mishkin, M. (2011). "A new neural framework for visuospatial processing." *Nature Reviews Neuroscience* 12, pp 217-230.

Kubilius, J., Wagemans, J., and Op de Beeck, H. P. (2014). "A conceptual framework of computations in mid-level vision." *Frontiers in Human Neuroscience* 8(158), pp 1-19.

Kveraga, K. and Bar, M. Eds. (2014). *Scene Vision: Making Sense of What We See.* Cambridge: MIT Press.

Landy, M. S., Maloney, L. T., Johnston, E. B., Young, M. (1995) "Measurement and Modeling of Depth Cue Combination: in Defense of Weak Fusion." *Vision Research* 35(3), pp 389-412.

Lu, Z., and Dosher, B. (2014). *Visual Psychophysics: From Laboratory to Theory.* Cambridge: MIT Press.

Lycan, W. (2015). "Representational Theories of Consciousness." *Stanford Encyclopedia of Philosophy.*

Mach, E. (1897). *The Analysis of Sensations.* New York: Dover, 1959.

Mach, E. (1902). "On the Psychology and Natural Development of Geometry." *The Monist* 12(4), pp 481-515.

Malcolm, G. L. and Schyns, P. G. (2014). "More Than Meets the Eye." In: Kveraga and Bar (2014), pp 27-44.

Maloney, L. T., Gerhard, H. E., Boyaci, H., and Doerschner, K. (2011). "Surface color perception and light field estimation in 3D scenes." In: Harris and Jenkin (2011), pp 280-307.

Marr, D. (1982). *Vision.* Cambridge: MIT Press, 2010.

Marr, D., and Hildreth, E. (1980). "Theory of Edge Detection." *Proceedings of the Royal Society of London Series B, Biological Sciences* 207(1167), pp 187-217.

Marr, D., and Poggio, T. (1977). "From Understanding Computation to Understanding Neural Circuitry." *Neurosciences Res. Prog. Bull.* 15(3), pp 470-488.

Marr, D., and Poggio, T. (1979). "A Computational Theory of Human Stereo Vision." *Proceedings of the Royal Society of London Series B, Biological Sciences* 204(1156), pp 301-328.

Marr, D. and Nishihara, K. (1978). "Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes." *Proceedings of the Royal Society B* 200(1140), pp 269-294.

Masrour, F. (2015). "The geometry of visual space and the nature of visual experience." *Philosophical Studies* 172, pp 1813-1832.

Masrour, F. (2016). "Space Perception, Visual Dissonance and the Fate of Standard Representationalism." *Noûs* 00(0), pp 1-29.

Matthen, M. (2005). *Seeing, Doing, and Knowing.* Oxford: Oxford University Press.

Matthen, M. (2010). "How things look (and what things look that way)". In B. Nanay (Ed.), *Perceiving the World.* New York: Oxford University Press.

Matthen, M. (2015). "Active Perception and the Representation of Space." In D. Stokes, M. Matthen, and S. Biggs. (2015). *Perception and Its Modalities.* Oxford: Oxford University Press.

Millar, S. (2008). *Space and Sense.* New York: Psychology Press.

Milner, A. and Goodale, M. (1995 2006). *The Visual Brain in Action*, 2nd edition. Oxford: Oxford University Press.

Moors, P., Boelens, D., van Overwalle, J., and Wagemans, J. (2016). "Scene Integration Without Awareness: No Conclusive Evidence for Processing Scene Congruency During Continuous Flash Suppression." *Psychological Science* 27(7), pp 945-956.

Moors, P., Wagemans, J., van Ee, R., and de-Wit, L. (2015). "No evidence for surface organization in Kanizsa configurations during continuous flash suppression." *Attention Perception Psychophysics* 79(3), pp 902-914.

Morgan, L. K., MacEvoy, S. P., Aguirre, G. K., and Epstein, R. A. (2011). "Distances between real-world locations are represented in the human hippocampus." *Journal of Neuroscience* 31(4), pp 1238-1245.

Nanay, B. (2017). "Philosophy of Perception: A Road Map with Lots of Bypass Roads." In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception.* New York: Routledge, pp 1-19.

Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., et al. (2011). "Scene-selective cortical regions in human and non-human primates." *Journal of Neuroscience* 31(39), pp 13771-13785 .

Newcombe, N., Huttenlocher, J., Sandberg, E., Lie, E., and Johnson, S. (1999). "What Do Misestimations and Asymmetries in Spatial Judgment Indicate About Spatial Representation?" *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25(4), pp 986-996.

Norman, J. F., Todd, J. T. and Phillips, F. (1995). "The Perception of Surface Orientation from Multiple Sources of Optical Information." *Perception and Psychophysics* 57, pp 629-636.

Oliva, A., and Torralba, A. (2001). "Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope." *International Journal of Computer Vision* 42(3), pp 145-175.

Oliva, A. (2005). "Gist of the Scene." In L. Itti et al. (Eds.), *Neurobiology of Attention.* New York: Academic Press, pp 251-257.

Oliva, A., Park, S., and Konkle, T. (2011). "Representing, perceiving, and remembering the shape of visual space." In Harris and Jenkin (2011), pp 308-336.

Oliva, A. (2014). "Scene Perception." In Werner and Chalupa (2014), pp 725-732.

Ooi, T. L., Wu, B. He, Z. J. (2001). "Distance determined by the angular declination below the horizon." *Nature* 414, pp 197-200.

O'Regan, K. (1992). "Solving the 'real' mysteries of visual perception: the world as an outside memory." *Can J Psychol.* 46(3), pp 461-488.

Orlandi, N. (2014). *The Innocent Eye.* Oxford: Oxford University Press.

O'Shaughnessy, B. (1980). *The Will: A Dual Aspect Theory, i.* Cambridge: Cambridge University Press.

Palmer, S. (1999). *Vision Science: From Photons to Phenomenology.* Cambridge: MIT Press.

Park, S., Brady, T. F., Greene, M., and 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." *Journal of Neuroscience* 31(4), pp 13331340.

Park, S. and Chun, M. M. (2014). "The constructive nature of scene perception." In Kveraga and Bar (2014), pp 45-72.

Park, S., Konkle, T., Oliva, A. (2015). "Parametric coding of the size and clutter of natural scenes in the human brain." *Cerebral Cortex* 25 (7), pp 1792-1805.

Parkes, L., J. Lund, A. Angelucci, J. Solomon, and M. Morgan. "Compulsory Averaging of Crowded Orientation Signals in Human Vision." *Nature Neuroscience* 4(7), pp 739-44.

Peacocke, C. (1992). *A Study of Concepts.* Cambridge: MIT Press.

Pelli, D. G., Palomares, M., and Majaj, N. J. (2004). "Crowding is unlike ordinary masking: Distinguishing feature integration from detection." *Journal of Vision* 4, pp 1136-1169

Potter, M. C. (1975). "Meaning in Visual Search." *Science* 187, pp 965-966.

Potter, M. C. (1993). "Very Short-Term Conceptual Memory." *Memory & Cognition* 21(2), pp 156-161.

Potter, M. C. (1999). "Understanding sentences and scenes: The role of Conceptual Short Term Memory." In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief stimuli.* Cambridge, MIT Press, pp 13-46

Potter, M. C., Wyble, B., Hagmann, C. E., McCourt, E. S. (2014). "Detecting Meaning in RSVP in 13 ms per picture." *Attention, Perception, and Psychophysics* 76(2), pp 270-279.

Putnam, H. (1963). "An Examination of Grnbaums Philosophy of Geometry." In H. Putnam (1975), *Mathematics, Matter, and Method.* Cambridge: Cambridge University Press, pp 93-129.

Raudies, F., and Hasselmo, M. E. (2012). "Modeling Boundary Vector Cell Firing Given Optic Flow as a Cue." *PLoS Computational Biology*, 8(6): e1002553.

Redish, A. D. (1999). *Beyond the Cognitive Map: From Place Cells to Episodic Memory.* Cambridge: MIT Press.

Rensink, R. A. (2000). "The Dynamic Representation of Scenes." *Visual Cognition* 7, pp 17-42.

Rensink, R. A. (2002). "Change Detection." *Annu. Rev. Psychol.* 53, pp 245-77.

Rogers, B. J., and R. Cagenello. (1989). "Disparity Curvature and the Perception of Three-Dimensional Curvature." *Nature* 339, pp 135-37.

Rosenholtz, R., J. Huang, and K. Ehinger. (2012). "Rethinking the Role of Top-down Attention in Vision: Effects Attributable to a Lossy Representation in Peripheral Vision." *Frontiers in Psychology* 3(1), pp 1-15.

Rosenholtz, R. (2015). "Texture Perception." In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization.* Oxford: Oxford University Press, pp 167-186.

Rosenholtz, R. (2016). "Capabilities and Limitations of Peripheral Vision." *Annual Review of Vision Science* 2, pp 437-57.

Schyns, P. and Oliva, A. (1994). "From Blobs to Boundary Edges." *Psychological Science* 5(4), pp 194-200.

Scott, N. M., Leuthold, A., Sera, M. D., and Georgopoulos, A. P. (2016). "Differential neural activity patterns for spatial relations in humans: a MEG study." *Exp Brain Research* 234, pp 429-441.

Sellars, W. (1968). *Science and Metaphysics: Variations on Kantian Themes.* London: Routledge & Kegan Paul.

Shepard, R. and Cooper, L. (Eds.) (1982). *Mental Images and Their Transformations.* Cambridge: MIT Press.

Shepard, R. and Metzler, D. (1972). "Mental Rotation of Three-dimensional Objects." *Science* 171, pp 701-703.

Siegel, S. (2006). "Direct Realism and Perceptual Consciousness." *Philosophy and Phenomenological Research* 73(2), pp 378-410.

Siegel, S. (2010). *The Contents of Visual Experience.* Oxford: Oxford University Press.

Siegel, S. (2017). *The Rationality of Perception.* Oxford: Oxford University Press.

Sperling, G. (1960). "The information available in brief visual presentations." *Psychological Monographs: General and Applied* 74, pp 1-29.

Stevens, Kent A. (1979). "Surface Perception from Local Analysis of Texture and Contour." MIT Artificial Intelligence Laboratory. Technical Report 512

Stone, J. V. (2011). *Vision and Brain: How We Perceive the World.* Cambridge: MIT Press.

Strawson P. (1959). *Individuals.* Routledge.

Strawson, P. (1966). *The Bounds of Sense: An Essay on Kants Critique of Pure Reason.* Routledge.

Suppes, P. (1977). "Is Visual Space Euclidean?" *Synthese* 35, pp 397-421.

Suppes, P. (1995). "Some Foundational Problems in the Theory of Visual Space." In D. Luce *et al.* (Eds.), *Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow Indow on his 70th Birthday.* Erlbaum, pp 37-45.

Thompson, B. (2010). "The Spatial Content of Experience." *Philosophy and Phenomenological Research* 81(1), pp 146-184.

Ungerleider, L., and Mishkin, M. (1982). "Two cortical visual systems." In M. Ingle, M. Goodale, and R. Masfield (Eds.), *Analysis of Visual Behavior*. Cambridge: MIT Press, pp 549-586.

Vishwanath, D. (2011). "Visual information in surface and depth perception: Reconciling pictures and reality." In L. Albertazzi, G. van Tonder, and D. Vishwanath (Eds.), *Perception Beyond Inference: The Information Content of Visual Processes.* Cambridge: MIT Press.

Vanmarcke, S., Mullin, C., Van der Hallen, R., Evers, K., Noens, I., Steyaert, J., and Wagemans, J. (2016). "In the Eye of the Beholder: Rapid Visual Perception of Real-Life Scenes by Young Adults with and Without ASD." *J Autism Dev Disord* 46, pp 2635-2652.

Vanmarcke, S. and Wagemans, J. (2015). "Rapid gist perception of meaningful real-life scenes: Exploring individual and gender differences in multiple categorization tasks." *i-Perception* 6, pp 1937.

Vishwanath, D. (2014). "Toward a new theory of stereopsis." *Psychological Review* 121, pp 151-178.

Wagemans, J., van Doorn, A. J., and Koenderink, J. J. (2011). "Measuring 3D point configurations in pictorial space." *i-Perception* 2, pp 77-111.

Wagner, M. (2006). *Geometries of Visual Space.* Erlbaum.

Walther, D. B., Caddigan, E., Fei-Fei, L., and Beck, D. M. (2009). "Natural scene categories revealed in distributed patterns of activity in the human brain." *Journal of Neuroscience* 29 (34), pp 10573-10581.

Ward, E. J., MacEvoy, S. P., Epstein, R. A. (2010). "Eye-centered encoding of visual space in scene-selective regions." *Journal of Vision* 10(14), pp 1-12.

Warren, P. A., & Rushton, S. K. (2007). "Perception of object trajectory: Parsing retinal motion into self and object movement components." *Journal of Vision* 7(11):2, pp 1-11.

Warrington, E., and Taylor, A. (1973). "The contribution of the right parietal lobe to object recognition." *Cortex* 9, pp 152-164.

Werner, J. S. and Chalupa, L. M. (2014). *The New Visual Neurosciences.* Cambridge: MIT Press.

Wolbers, T., Klatzky, R. L., Loomis, J. M., Wutte, M. G., and Giudice, N. A. (2011). "Modality-independent coding of spatial layout in the human brain." *Current Biology* 21(11), pp 984-989.

Wu, B., Ooi, T. L., and He, Z. J. (2004.) "Perceiving Distance Accurately by a Directional Process of Integrating Ground Information." *Nature* 428(4), pp 73-77.

Wu, W. (2013). "Visual spatial constancy and modularity: Does intention penetrate vision?" *Philosophical Studies* 165, pp 647-669.

Wu, W. (2014). "Against Division: Consciousness, Information and the Visual Streams." *Mind & Language* 29(4), pp 383-406.

Wurtz, R. H., Joiner, W. M., and Berman, R. A. (2011). "Neuronal mechanisms for visual stability: progress and problems." *Phil. Trans. R. Soc. B* 366, pp 492-503.

Chapter 2

Are Euclidean Beliefs Grounded in Core Cognitive Systems?

2.1 Introduction

Over the past decade, philosophers have drawn upon psychological theories of numerical concept acquisition and arithmetical belief formation to defend claims about the metaphysics,semantics, and epistemology of arithmetic.[1] Most commonly, these philosophers have focused on one theory from the psychological literature that has come to be known as the "theory of core cognition." The theory of core cognition offers an account of concept acquisition and belief formation in at least five domains: *portable objects, intention action, number, geometry,* and *social group* (Spelke and Kinzler 2007). For each domain, a collection of unlearned core cognitive systems makes possible the acquisition of basic concepts and the formation of basic beliefs about this domain. Until recently, the domains of *object* and *number* have received the most attention from core cognitive theorists and philosophers of arithmetic. However, over the past decade, Spelke and others have published a number of articles developing the core cognitive theory of *geometry.* The primary purpose of this essay

[1] For example, Ball (2017), Beck (2015), Burge (2010, pp 471ff), Giaquinto (2001a,b), Heck (2000), Jenkins (2005, 2008), Maddy (2007, 2014), Margolis and Laurence (2008), Prinz (2002, pp 184ff), and Shea (2011). Other philosophers of arithmetic, such as Hale and Wright (2001), Parsons (2009), and Resnik (1997), defend their metaphysical, semantic, and epistemological claims without engaging this experimental literature.

is to evaluate whether the theory of core cognition might be as useful for philosophers of geometry as it has become for philosophers of arithmetic. My answer will be largely negative given the current state of scientific knowledge.

In Part 1, I review the core cognitive theory of numerical concept acquisition in order to isolate the essential claims of the core cognitive theory of number. In Part 2, I show how core cognitive theorists have defended parallel claims in the domain of geometry. In Part 3, I evaluate the strength of the core cognitive theory and reflect on its utility for philosophers of geometry.

2.2 The Core Cognitive Theory of Numerical Concept Acquisition

In her *Origin of Concepts*, Susan Carey proposes a general theory of concept acquisition. According to this theory, the content of a learned concept in one of the five domains (e.g., the numerical concept *eight*) is determined by its causal relations with objects in the environment. These causal relations are sustained by a set of unlearned, domain-specific, modular systems of representation: core cognitive systems. The representations in these core cognitive systems are necessarily intermodal and sensory-derived (Carey 2009, pp 38-40, 136). Additionally, core cognitive representations are non-conceptual in the sense that they cannot contain logical constants (e.g., negation or disjunction), cannot be asserted, and cannot be entertained without attributing the property it represents to a particular in the environment.[2] According to Carey, concepts *cannot* be acquired by combining the representational contents of core cognitive systems.[3] Rather, a set of interrelated concepts (e.g., natural number concepts) is acquired when a corresponding set of lexical items (e.g.,

[2]Carey claims that core cognitive representations are "conceptual," but by this she merely means that they are post-perceptual (2009, p 34) and intermodal (2009, p 39). I believe Carey would agree that core cognitive representations lack all three aspects of my more restricted definition of *concept*.

[3]To claim that concepts are acquired in this way is to fall prey to Fodors (1975, p 36ff) *reductio* in favor of radical concept nativism (Carey 2009, p 18).

numerals) are appropriately related to each other and to the underlying core cognitive systems. Carey draws this theory of concept acquisition from philosophical essays like Laurence and Margolis (1999, 2002) and she attributes it to many psychological essays from the early 1990s.[4] Nevertheless, her book is currently the most detailed and experimentally grounded explication of the core cognitive theory of concept acquisition.

Carey applies the core cognitive theory of concept acquisition to the system of natural number concepts (2009, ch 4). In particular, she wants to give an account of "how children learn the meaning of verbal numerals such as 'three' and 'seven' " (p 134). Possession of this system of concepts entails that one can correctly answer questions about the cardinality of sets of objects, such as whether two sets have equal cardinality or questions of how many objects are in the union or difference of two sets. Carey claims that the causal relations determining the contents and acquisition of numerical concepts are sustained by two core cognitive systems: the analogue magnitude system (AM) and the parallel individuation system (PI).[5]

2.2.1 Two Core Systems of Number

The AM system is responsible for representing the numerical size of a set of environmental particular. Core theorists individuate the AM system by its conformity to Weber's law.

[4]Carey (2009, pp 25, 306, 307, 511, 517) also cites MacNamara (1986), Block (1986), and Nersessian (1992) as philosophical articulations of the bootstrapping process that is espoused in the theory of core cognition. Carey maintains that the core cognitive theory of *numerical* concept acquisition was first defended in five studies from the early 1990s (p 25n2), and we might add to her list the publications of Fuson (1988), Wynn (1990), and Dehaene and Changeux (1993). Despite some similarities, the core cognitive theory of concept acquisition should not be attributed to Piaget. Piaget (1936) falsely claims that certain object and number representations do not develop until around age 5, but these representations have since been isolated in six-month-old infants (Carey 2009, pp 121, 136). Still, the core cognitive theory and Piagets theory both posit discrete stages of cognitive growth and concept acquisition. This distinguishes their paradigms from more 'empiricist,' 'gradualist,' or 'non-modular architecturalist' approaches, such as Bates *et al.* (1998), Karmiloff-Smith (1992), Spencer et al. (2009), and Thelen and Smith (1994).

[5]Carey (2009), pp 117, 137. The AM system is identical to what Dehaene (2011) calls the "analogue number sense" (Carey 2009, p 118). The PI system tracks multiple individuals (e.g., objects, events, tone bursts) in working memory and during periods of occlusion (p 117).

Weber's law maintains that the discriminability of the (numerical) size of any two sets is a function of the ratio of their cardinalities (2009, p 118). For example, according to Weber's law, it should be more difficult for an animal to determine whether an 8-object set and a 9-object set have the same size than it is to determine whether an 8-object set and a 16-object set have the same size. Many species, including humans, respond to cardinality properties in accordance with Webers law. For example, Elizabeth Brannon and colleagues have shown that Rhesus monkeys can learn to order a novel collection of images by the number of closed regions in each image.[6] The Rhesus monkeys were successful at ordering the images by cardinality when every pair of images had a cardinality ratio greater than 3:2, but they failed when the ratio was smaller than this. Their performance conforms to Webers law. In a second study, Xu and Spelke (2000) have shown that six-month-old human infants *dishabituate* to (i.e., look longer at) a new dot array if its cardinality ratio to the old dot array is greater 3:2.[7]

The AM system is not limited to representing that the numerical size of one set is not equal to (or greater than) the numerical size of another set. The AM system can also represent and compute the addition of two numerical sizes. Flombaum *et al.* (2005) showed that Rhesus monkeys dishabituate if two sets of n objects are put behind a screen and then the screen is lifted to reveal n objects rather than $2n$ objects.[8] Still, it would be inaccurate to say the AM system contains *arithmetical* representations. There is no evidence that AM representations contain singular reference to discrete natural numbers (e.g., "the number of planets is eight", "seven is the successor of six").[9] Rather, the Weber-fraction signature suggests the AM system represents 9 numerical size as a continuous property similar to

[6]Brannon and Terrace (1998), Cantlon and Brannon (2006).

[7]Spelke replicated this result using tone sequence cardinalities (Lipton and Spelke 2003, 2004) and cardinalities of jumping event sequences (Wood and Spelke 2005).

[8]This experiment is replicated for nine-month-old infants in McCrink and Wynn (2004).

[9]As Carey notes, an outdated model of the AM system called the "accumulator model" did posit such discrete representations that are computed via a subconscious counting routine (2009, p 293). But the results of Barth *et al.* (2005) and Wood and Spelke (2005) show that the time it takes adults and infants to form an AM representation is invariant across set size.

perceptible magnitudes like brightness, length, and temperature. Before we can explain the acquisition of natural number concepts, we must consider the representational contents of the PI system.

Core theorists individuate the PI system and their representations by their conformity to a "set-size signature" (p 139). This signature arises in habituation studies where objects are occluded one at a time behind a screen or in a jar. Starkey and Cooper (1980) found that, although infants dishabituate when two objects are hidden and three are revealed, infants do *not* dishabituate when four objects are hidden and six are revealed. These results violate Weber's law. Hence, the AM system is not responsible for the infant's ability to discriminate the sizes of these sets. Core cognitive theorists claims that the infants discrimination abilities are grounded in an object-tracking system that is incapable of tracking more than three objects. Their hypothesis is confirmed by another experimental paradigm involving jars of occluded food items. In this paradigm, each of two empty jars are visibly and sequentially filled with a certain number of graham crackers (Feigenson and Carey 2005). Infants are then allowed to grab the food out of one of the jars. Infants consistently choose the numerically larger set if both jars contain fewer than three objects. However, if either jar contains four or more objects, they choose at random between the two jars (even in the case of one versus four!).

There is some controversy about how we should characterize the content of the PI system. Some core theorists believe that the above experiments show that the PI system contains applied arithmetical representations of, e.g., $2 + 2 = 4$. Other core theorists believe that the PI system merely opens and closes individual-files rather than attributing cardinality to sets.[10] Despite this controversy, it is universally accepted that the PI system is not capable of representing arithmetical facts involving numbers greater than four. The core cognitive theory of number maintains that applied representations of, e.g., $5 + 7 = 12$, are only possible

[10]Carey (2009, pp 142-143) seems to hold a position between these two extremes according to which infants can represent that two sets stand in a 1-1 correspondence.

through the childs construction of the verbal numeral list.[11] As we will see below, the PI and AM core systems play "a crucial role in the creation of the explicit verbal numeral list representation of the positive integers" (Carey 2009, p 152).

2.2.2 Bootstrapping to Numerical Concepts and Arithmetical Beliefs

In any domain, the theory of core cognition desires to show how it is possible for language learning to "build representational resources that transcend core cognition" (p 248). In any domain, this process requires acquiring a lexicon along with an associated network of inference and action rules (p 306). In the domain of number, core theorists agree that children acquire numerical concepts by first learning: (i) to rehearse the verbal numeral list (*one, two, three,* etc.), (ii) to count a set of objects, and (iii) to answer 'how-many' questions with the the last numeral uttered during the counting procedure.[12] Tokens of these numerals lack meaning until they can be used to correctly answer a variety of cardinality questions about sets of environmental particulars.[13]

What allows a child to correctly answer cardinality questions? Core theorists believe that the singular/plural distinction in English grammar is the primary catalyst for correctly answering these questions. They claim that children learn the meaning of the word "some" as appropriate in cases where a visible set has two or more elements. They also claim that many experimental results show that when children start to correctly answer cardinality questions for sets with one element, the are actually mapping the numeral one to the indefinite article

[11]We may assume that core theorists would also accept that non-verbal numeral lists can also be used a basis for the development of arithmetical representations.

[12]Carey calls these three skills, "the stable order principle," "the 1-1 correspondence principle," and "the last word rule," respectively (2009, pp 288-290, 300). As many experimental results show, these skills are not sufficient for numerical concept possession since children with these skills are unsuccessful on the three tasks described in the next footnote.

[13]These questions include: "Can you give me n?" (Give-a-Number task), "Can you point to the card with n fish?" (Point-to-n task), and "What's on this card?" (What's-on-this-card task). When a child is successful at all of these tasks for any set size, core theorists say they have acquired the "cardinality principle." Cf. Carey (2009), pp 297-302.

"a" and all numerals greater than one to the word "some" (p 326). The crucial step comes when they learn to correctly answer cardinality questions for sets of two, three, and four elements. In the case of "two," children learn to apply this numeral just in case their PI system contains two open individual files. In every such case, the numeral "two" functions as a determiner and thus describes a property of a *set* of objects. This property is naturally inferred to be the numerical size that is represented in the AM system whenever the PI system has two open individual files. In this way, "two" comes to represent a numerical size. The same process explains how the numerals "three" and "four" come to represent numerical sizes. In all three cases, when a child has learned to utter the numeral n whenever they have n open individual files, their utterance will have the same meaning as their AM representation of that set.

The PI system is not capable of opening more than four individual files. For this reason, higher numerals only gain reference to a precise numerical size through a childs ability to enumerate a set. The child's disposition to answer 'how-many' questions using the last word rule combines with singular/plural distinction to activate an AM representation of the enumerated set. Through this procedure, higher numerals gain their reference to numerical size. These numerals still do not represent cardinal numbers because they "do not yet embody the arithmetic successor function" (p 327). However, children can acquire the successor function concept when, for any set of size n, the meaning of numeral $n + 1$ is simply the size of the set resulting from adding a new individual" (p 327). Children then use "this integrated representational system to invent addition and subtraction algorithms based on the successor function (i.e., by counting up and counting down)" (p 344). To summarize, the singular/plural distinction in English grammar makes possible the simultaneous activation of the PI and AM systems, which in turn allows the network of inference rules (i)-(iii) to associate numerals with discretized content.

When children are able to correctly answer cardinality questions, the system of numerical concepts they now possess is constitutively related to core representations in the AM and PI systems. That is, the representations in these two core systems are capable of producing basic arithmetical beliefs. For example, if adults are shown two dot arrays where the first array has twice as many dots as the other array, the AM system will represent this fact and give rise to a belief *that the number of dots in the first array is greater than the number of dots in the second array* (Barth *et al.* 2003). When the number of dots in the two arrays is varied, Barth and colleagues have shown that the accuracy of adult verbal reports about which array is larger conforms to Webers law.[14] The mechanism underlying this sort of basic arithmetical belief formation is sustained by "a mapping between analog magnitude representations and the verbal integer list" (Carey 2009, p 128). Similarly, if the PI system has three object files open, we are capable of forming a belief *without counting* that there are three objects in the visible set.[15] In this way, the theory of core cognition is just as much a theory of belief formation as it is a theory of concept acquisition.

In order to evaluate the core cognitive theory of geometry, I want to isolate the essential claims of the core cognitive theory of number. There are three such claims: (a) the AM and PI systems exist, (b) numerical concept acquisition occurs when these two core systems allow the numeral list to be appropriately applied to the environment, and (c) the representations in the AM and PI system dispose us to form corresponding basic beliefs. Many philosophers of arithmetic have adopted at least one of these claims in developing their own metaphysical, semantic, and epistemological arguments.[16] They have used this theory and its supporting

[14]Carey (2009), p 131. Barth *et al.* (2006) also showed human adults and four-year-old children a serial display of a blue 45-dot array, a blue 60-dot array, and a red 75-dot array. They asked participants whether they saw more blue dots or red dots. Their performance was again predicted by Weber's law, which suggests that the AM system generates arithmetical beliefs involving addition-like operations. (Carey 2009, pp 123, 127, 128)

[15]The exercise of this capacity is often called "subitizing" (Dehaene and Cohen 1994).

[16]Philosophers who support parts of the core cognitive theory of numerical concept acquisition include Ball (2017), Beck (2015), Maddy (2007, 2014), Shea (2011), Burge (2010), Laurence and Margolis (2008), and Giaquinto (2001). Some of these endorsements come with dissenting claims about some aspect of Carey's view of the representations in these two core cognitive systems: Ball (2017) argues they are not iconic, Beck (2015) argues they are not conceptual, Burge (2010, p 480n82) argues they do not represent numerosity. I

experiments to refine and extend their own theories of, e.g., arithmetical belief formation, knowledge, analyticity, apriority, truth, and meaning. It is only natural to ask whether core cognitive theory can bear philosophical fruitful in other domains. In particular, I will ask whether the core cognitive theory of geometrical concept acquisition and belief formation can stimulate theorizing in the philosophy of geometry.

2.3 The Core Cognitive Theory of Geometric Concept Acquisition

According to Spelke *et al.* (2010), core cognition of geometry parallels the structure of core cognition of number. There are two core systems – the *surface layout system* and the *object shape system* – that make the acquisition of familiar geometric concepts and Euclidean beliefs possible:

> What, then, is the nature of human knowledge of geometry, and how does this knowledge arise and develop? Here, we offer a hypothesis in the spirit of Carey (2009) ... Like natural number, natural geometry is founded on at least two evolutionarily ancient, early developing, and cross-culturally universal cognitive systems that capture abstract information about the shape of the surrounding world: two core systems of geometry. Nevertheless, each system is limited: It captures only a subset of the properties encompassed by Euclidean geometry, and it applies only to a subset of the perceptible entities to which human adults give shape descriptions. Children go beyond these limits, and construct a new system of geometric representation that is more complete and general, by combining productively the representations delivered by these two systems. (2010, p 865)

agree with these claims, but they should not be seen as challenging the essential claims of the core cognitive theory of numerical concept acquisition.

The "new system of geometric representation" refers to a collection of propositions and lexical concepts found in Euclid's Elements. In the case of arithmetic, the verbal numeral list and the singular/plural distinction made the acquisition of numerical concepts and arithmetical beliefs possible. In the case of geometry, acquisition is made possible by a set of "uniquely human, culturally variable artifacts: pictures, models and maps" (2010, p 865).[17] These artifacts make possible the construction of a rich relational structure (exhibiting typical Euclidean relations between points, lines, and angles) that is responsible for producing basic Euclidean beliefs.[18] This rich relational structure is thought to augment the geometric content of cognitive maps, working memory, and the spatial representations inferred from drawn maps.[19] In short, the core cognitive theory of geometry posits three levels of representation: the two core systems, an artifact-based rich relational structure, and Euclidean concepts and beliefs. Over the past decade, articles by Spelke and many other cognitive psychologists have argued for the existence of these representational systems and various causal dependencies between them on the basis of behavioral and neural evidence that I will describe below.[20]

[17]Spelke *et al.* (2010) make this point explicitly: "Thus, like the system of number, the system of geometry that feels most natural to educated adults is a hard-won cognitive achievement, constructed by children as they engage with the symbol systems of their culture" (p 865).

[18]For example, Dillon and Spelke (2018, p 1) claim that "older childrens map reading undergoes changes through development that predict the emergence of intuitions that align better with euclidean geometry."

[19]Dehaene *et al.* (2006) make the same point: "Geometrical intuitions, in the final analysis, may rest on a spontaneous imposition of stable conceptual relations onto variable and imperfect sensory data" (p 383). There is some disagreement about whether all three of these spatial representations take on this Euclidean structure. For example, Lee *et al.* (2012, p 145) claim that certain experimental results demonstrate that cognitive maps do not have a Euclidean structure. Rather, the geometric content of cognitive maps is closer to what Chrastil and Warren (2014) refer to as "graph knowledge."

[20]Calero *et al.* (2019), Dillon and Spelke (2015, 2018), Dillon *et al.* (2013), Lee *et al.* (2012), Spelke and Lee (2012), Izard *et al.* (2011a), Spelke (2011), Spelke *et al.* (2010), Izard and Spelke (2009), Dehaene *et al.* (2006). Within broadly the same tradition, some cognitive psychologists (e.g., Landau 2017, Shusterman and Li 2016) have focused on the acquisition of non-metric concepts expressed by locative prepositional phrases (e.g., 'on', 'to left of', 'uphill from'). Still other articles in this tradition (e.g., Winkler-Rhoades *et al.* 2013, Burgess 2006) have merely defended the view that the two core systems of geometry can explain the origins of other cognitive, but pre-conceptual, skills such as the construction and application of cognitive maps.

2.3.1 Two Core Systems of Geometry

The core cognitive theory of geometric concept acquisition appeals to a variety of neural and behavioral evidence to argue for the existence, neural basis, and content of the surface layout system and the object shape system.[21] I will give an overview of this evidence here. Let's begin with the surface layout system. This core system, also called the "geometric module," was first proposed by Ken Cheng (1986) based on the results of three experiments about the reorientation behavior by rats. In each experiment, rats were trained to find hidden food at up to nine locations in a rectangular box. In the test trials, after rats learned the location of the food, they were removed and placed in an identical box where they had to reorient to find the food again. Cheng found that features such as the color, texture, and lighting on the walls were *not* used to reorient. Based on these results, he posits a core system in the working memory of rats that guides their navigation to a target object in a familiar environment. In this core system, Cheng claims that rats attribute a "metric frame" to the environment based on the "arrangement of surfaces," and they location food *via* an "address on the metric frame" (p 172).

A decade after Cheng's experiments, neuroscientists recorded the activity of hippocampal "place cells" while rats explored a rectangular environment (O'Keefe and Burgess 1996). They found that the firing rates of certain place cells varied with the rat's distance from a given boundary wall. Further studies compared the firing rates of specific place cells across variously-shaped environments, and they concluded that "the representation of space provided by place-cell firing rates is like a single flexible map, or more accurately a coordinate system, which can be applied to any environment" (Hartley *et al.* 2000, p 378). By this, they mean that place cell firing rates are computed from a set of postulated "boundary vector cells" that each represent the distance to a boundary wall (and thus can be thought of as one coordinate in the metric frame). Boundary vector cells were isolated in the rats entorhinal

[21] Unlike the numerical case, these two systems have not been given widely-adopted proper names.

cortex by Solstad *et al.* (2008).[22] In the case of humans, the fMRI studies of Doeller *et al.* (2008) have found that when a human is being taught the location of a hidden object (relative to a circular wall and the sun), the activity levels in the right posterior hippocampus predict how well the subject is able to estimate that object's location in a future trial. Additionally, Hupbach and Nadel (2005) and Spelke and Lee (2008) found that two-year-old children do *not* reorient on the basis of the distance between landmarks (e.g., a pole) or the size of the angle between boundary walls.[23]

Core cognition theorists, who are primarily concerned with concept acquisition, regularly cite the above studies to elaborate the contents and neural basis of the spatial layout system. For example, Spelke *et al.* (2010, pp 867ff) claim that these studies demonstrate that the spatial layout system represents (i) the egocentric distance and direction to each boundary wall, (ii) the proportion between the lengths of adjacent boundary walls,[24] and (iii) the left/right relations between pairs of adjacent boundary walls. In contrast, these representations do not include contents such as the distance between landmarks, the *color*

[22]This study is often cited by core theorists like Spelke and Lee (2012, p 2786), Lee *et al.* (2012), p 157. In addition to these boundary vectors cells, other kinds of newly-isolated cells have been implicated in navigation. For instance, Hafting *et al.* (2005) isolated "grid cells." If we imagine the navigable environment has been tiled by equilateral triangles, each of these grid cells fires when the location of the animal has moved over a vertex of a triangle. They propose that this firing pattern is evidence of a metric space representation of the environment (often called a *cognitive map*), in which the animal stores the location of objects, landmarks, and boundaries.

[23]As recently as Lee *et al.* (2012, pp 153, 157), core theorists have appealed to boundary vector cells as the neural basis of the spatial layout system. However, there have been more recent attempts (e.g., Dillon *et al.* 2017) by core theorists to argue that the neural basis of the spatial layout system is in fact the occipital place area (OPA) and the retrosplenial complex (RSC). The argument that these areas represent distance and left/right relations is based on fMRI adaptation results showing that these two brain regions *release* from adaptation when the distance to a pictured landmark changes or the image is mirror-reversed. Unfortunately, these results about adaptation release can be explain by any number of proximal image features or non-geometric distal properties. Moreover, the OPA (and, arguably, the RSC) is a part of the *visuo-perceptual system*, and hence cannot be the neural basis of a core system. For these reasons, in the rest of this essay I will assume that boundary vector cells are the intended neural basis of the spatial layout system.

[24]Spelke *et al.* (2010), p 868. Core theorists seem to have rejected claim (ii) after Lee *et al.* (2012, experiments 7 and 8) showed that when four disconnected walls of equal length were arranged in a rectangular shape, children were able to reorient. However, when four disconnected walls of two different lengths were arranged in a square shape, children failed to reorient. This strikes me as a good reason to drop claim (ii). If we do drop this claim, none of my critiques of the core cognitive theory of geometry would be affected, but the diagram below from Spelke *et al.* (2010) would need to be updated.

of, *texture of*, or *angle between* these boundary walls. The former set of properties, or so it is maintained, innately and spontaneously guide human reorientation behavior in familiar environments, while the latter properties do not.[25] Since this system does not represent certain geometric features (e.g., angles and distance) it "fails as a system of Euclidean geometry" (Spelke *et al.* 2010, p 867).

The object shape system also fails as a system of Euclidean geometry but for slightly different reasons. The object shape system represents the geometric relations between the parts of 3D portable objects (rather than between navigational boundaries). Spelke *et al.* (2010) claim that this system represents (i) the size of angles formed by two edges of the object, and (ii) the proportion between the length of adjacent edges. Notably, this system does *not* represent the left/right relations between edges. The main results that core cognitive theorists use to support these claims come from *oddity tasks*. In an oddity task, participants are shown six roughly-similar figures simultaneously and asked to select the 'odd' figure. Dehaene *et al.* (2006) found that pre-adolescent children in the U.S. do well at selecting the figure with distinct angles or distinct length proportions. They also found that in trials where the odd figure is a mirror image of the other figures (at various rotations), children only make the correct selection 23% of the time (p 382). According to Spelke *et al.* (2010), the neural basis of this core system is identical to the neural basis of visual shape perception: the lateral occipital cortex (LOC). Dilks *et al.* (2011) have shown that the LOC exhibits adaptation across serial displays of mirror-reversed figures. This adaptation effect suggests that visual shape perception does not represent the left/right relations between the edges of the perceived figure. Based on these findings, core cognitive theorists often claim that the object shape system represents angles and length proportions, but not left/right ("sense") relations.

[25]The non-geometric properties are allegedly capable of guiding reorientation behavior, but only by means of non-modular "associative learning" (Doeller and Burgess 2008; Spelke *et al.* 2012).

2.3.2 Bootstrapping to Geometric Concepts and Beliefs

According to the theory of core cognition, the two core systems of geometry make possible the acquisition of geometric concepts and Euclidean beliefs.[26] The two core systems of geometry are intended to be structurally parallel to the two core systems of number. Spelke *et al.* (2010, p 874) make this parallel explicit in their Figure 4. In the case of geometric concept acquisition, the two core systems of geometry allegedly sustain the causal relations that give meaning to the relevant lexicon (e.g., *point, line, angle, circle, triangle, parallel*). When an adolescent human has acquired these geometric concepts, Spelke *et al.* (2010, p 878) describe this achievement as the "construction of a more general, unified system of Euclidean geometry." Dillon and Spelke (2018) operationalize the unity of this Euclidean system by means of (i) a triangle completion task, and (ii) an angle-based reorientation task.[27] Both tasks require the participant to infer the size of one angle in a triangle (or a rhombus) from core representations of side lengths. Two-year-old children perform no better than chance on these two tasks, but twelve year-olds are successful at both tasks. What allows an adolescent to successfully complete these tasks? As in the case of number, certain cultural artifacts are supposed to explain this bootstrapping process. In core number cognition, these artifacts were the verbal numeral list and the singular/plural distinction. In core geometric cognition, these artifacts are supposed to be pictures, models, and (primarily) maps. Similar to the singular/plural distinction, maps can simultaneously activate both core systems. To take one example, an overhead line map of a triangular space activates the object shape system because it is itself a small portable object with a 2D shape *and* it activates the spatial layout system because it represents a large, navigable space with boundary walls. As of yet, the core

[26] Carey (2009) labels concept acquisition processes, "bootstrapping," when they constitutively rely on core systems for their representational content.

[27] In this study, the *triangle completion task* required the participants to predict whether the (imagined) third angle in a triangle would get *smaller* or *larger* based on seeing the base angles either (i) increase in size or (ii) move farther apart (Dillon and Spelke 2018, p 8, fig 3). The *angle-based reorientation task* provided a map of either three disconnected sides or three disconnected angles that the participant had to complete in mental imagery in order to reorient in a disconnected enclosure of the opposite kind (2018, p 5, fig 1).

cognitive explanation for how this simultaneous activation is supposed to confer meaning on geometric concepts like *triangle*, *length*, and *parallel* is not as detailed in the geometric case as it is in the arithmetical case.

The core cognitive framework for geometric concept acquisition and Euclidean belief formation is appealing. In time, something like it may prove to be correct. However, I want to show here that the theory is inadequate as it currently stands. There are problems with the claims about the two systems, the claims about the resulting concepts, and the claims about the basic Euclidean beliefs that the core systems allegedly produce. These problems should serve as a cautionary tale for any philosophers trying to building philosophical arguments about truth, meaning, and knowledge in geometry while at the same time assuming the accuracy of the core cognitive theory of geometric concept acquisition.

2.4 Problems for the Core Cognitive Theory of Geometry

The core cognitive theory of geometric concept acquisition faces three major obstacles. First, the two core systems lack clear empirical support in the face of more recent neural and behavioral evidence. Second, the geometric concepts that children allegedly acquire have proven much more difficult to isolate than the analogous numerical concepts. Third, the Euclidean beliefs that these core systems are supposed to generate have not been directly tested.

2.4.1 Critique of the Two Core Systems of Geometry

The theory of core cognition posits the existence of core systems. By definition, core systems are post-perceptual and intermodal. The *object shape system* that Spelke and other core theorists posit, however, is a visuo-perceptual system. The neural basis of this system, the lateral occipital cortex, is squarely within the vision-specific ventral stream that is responsible for visual perception (Kanwisher and Dilk 2014, p 733). Hence, it is clear that the object shape system cannot be a core system. It is certainly possible there is an intermodal core system with similar geometric contents to the object shape system (i.e., *angles*, *length proportions*, but not *left/right relations*). But the neural and behavioral evidence that Spelke and others have cited does not support this claim.

The *spatial layout system* fairs a bit better. This core system has a neural basis in the boundary vector cells of the hippocampus. As such, it is both post-perceptual and intermodal. Nevertheless, there are independent reasons for doubting that this system can play the bootstrapping role that core theorists want it to play. The first reason is that these boundary vectors cells in the hippocampus are part of a "a wider network of spatially modulated neurons, including grid, [place], and head direction cells, each with distinct roles in the representation of space and spatial memory" (Moser *et al.* 2015, p 2). This wider network of spatially modulated neurons is not limited in the same way that boundary cells are limited. In particular, the firing patterns of grid cells and place cells reveal that rats represent their self-location on a metrically-specified grid of the navigable environment (Moser *et al.* 2015, p 3). Surprisingly, the firing patterns of the grid cells form equilateral triangles that tile the surrounding space (2015, p 3, fig 1). The precise geometric contents of this wider network of spatially modulated neurons is not obvious. However, it is clear that the failure of boundary vector cells to represent angle sizes does not adequately account for the contents of this wider network. Spelke and other core theorists have not provide any reason to think that the spatial layout system should exclude the richer geometric contents of place, grid, and head direction cells. Until such reasons are given, the neural evidence cited by core theorists in favor of a spatial layout system does little to support their claim that this system cannot represent the size of angles.

The second reason for doubting the existence of the spatial layout system comes from more recent reorientation studies. As we saw above, the core theorists identify the spatial layout system with Ken Chengs (1986) proposed "geometric module." His experiments were taken to show that wall color, brightness, and texture were not used by rats to reorient. Later, core theorists used Cheng's experimental set up to show that these claims apply to two-year-old children as well. (Further results about the lack of angle-based reorientation came from these later studies as well.) But, notably, Cheng (2008 p 355) now maintains the behavioral evidence for the existence of his geometric module "crumbled" beginning with two reorientation studies published around 2005. These studies demonstrated that rats and chickadees reorient in a rectangular enclosure based on the *color* of walls, which overshadow the geometric properties of these walls.[28] This evidence led Cheng to conclude that the system underlying reorientation is associative and image-based, rather than modular or geometric:

> [New empirical] developments have led to an associative learning theory modeling how featural cues [like color and texture] can sometimes help and sometimes hinder the learning of geometric cues. And they have led to an approach of view-based matching, in which geometric properties are not explicitly encoded. (2008, p 355)

The behavioral evidence against the geometric module has only grown since 2008.[29] Cheng and his colleagues often explicitly note that these new studies undermine the core cognitive theory of geometry (Twyman and Newcombe 2010, Sutton and Newcombe 2014). In fact, Newcombe *et al.* (2009) use these newer studies of reorientation as evidence all modular theories of mind. I do not mean to endorse these claims, but merely to note that the spatial layout system lacks the clear behavioral support that undergirds the two core systems of

[28]These studies are Gray *et al.* (2005) and Pearce *et al.* (2006).

[29]Twyman *et al.* (2013) and Cheng *et al.* (2013) provide a nice overview of the many studies disconfirming the existence of a geometric module.

arithmetic. More generally, the neural and behavioral results that initially seemed to support the object shape system and the spatial layout system have proved difficult to square with more recent results.

2.4.2 Critique of the Proposed Geometric Concepts and Beliefs

The core cognitive theory of geometry is a theory of how we acquire geometric concepts mentioned in Euclids *Elements* (e.g., *point*, *line*, *equilateral triangle*). This lexicon is notably dissimilar to the lexicon of numerals that is studied by core numerical cognition. For one, what set of concepts must be learned in order to have genuinely "Euclidean" beliefs? Are concepts like *left*, *right*, *rhombus*, and *cube* necessarily in the set of geometric concepts whose acquisition core theorists are trying to explain? In the case of number, there is a relatively precise set of numerical concepts, function concepts, and relation concepts (e.g., *greater than*, *addition of*, *number of*, *successor of*) that delimit the constituents of arithmetical beliefs. In the case of geometry, I am not aware of any similarly precise proposals for a set of basic geometric concepts. The only statement of which I am aware that describes the intended basic lexicon of geometry claims that it is composed of two syntactic kinds: *shape names* and *locative prepositions* (Spelke *et al.* 2010, p 874). Without a more precise sense of these set of basic geometric concepts are required, it will be difficult to determine when children typically acquire these concepts and to theorize which core systems may be involved in their acquisition. In the meantime, Dillon and Spelke (2018) use the Peabody Picture Vocabulary Test (PPVT) to gauge mastery of geometry. They show that the PPVT score predicts success on the triangle completion task and the angle-based reorientation task. However, given that the PPVT is primarily a test of everyday sortal concepts, these results tell us little about how basic geometric concepts are acquired. Hence, until a more precise lexicon is established, it will be difficult to develop an experimental paradigm that can provide support for the core cognitive theory.

Perhaps the more fundamental problem is that no precise set of Euclidean *beliefs* has been experimentally established. Although the propositions of Euclid's *Elements* are often cited as the set of Euclidean beliefs, the experimental procedures do not test for such beliefs. In discussions of both the triangle completion task and the angle-based reorientation task, core theorists often claim that successful participants must thereby believe a Euclidean proposition. In the triangle completion task, it is that the sum of the three angles in a triangle is a constant. In the angle-based reorientation task, it is that length of the sides determine the size of the angles in a triangle. However, this is questionable. It seems clear that participants would be able to successfully complete either task on the basis of amodal completion or visual image generation. These latter abilities are not tied to Euclidean belief formation in any obvious way.[30] The one task that core theorists have used in their experiments that seems to have the potential to reveal the formation of Euclidean beliefs is the line intuitions task of Izard *et al.* (2011b). In this task, participants are shown diagrams of, say, three non-collinear points and asked questions like, "Can a straight line be drawn through all three points?" Other questions in the line intuitions task are about intersection and parallelity. Unfortunately for core theorists, the sample size is small and the results suggest that adults and children perform equally well on these questions. In sum, the core cognitive theory of geometric concept acquisition has faced unique difficulties in delimiting its explanandum and in providing evidence for the existence of its explanans. These problems should serve as a cautionary tale for any philosophers trying to build philosophical arguments about truth, meaning, and knowledge in geometry while at the same time assuming the accuracy of the core cognitive theory of geometric concept acquisition.

[30]There are some suggestive remarks from core theorists that these visual imagery abilities are dependent on the formation of Euclidean beliefs. However, I am not aware of any evidence for this counterintuitive claim.

2.5 Bibliography

Ball, B. (2017). "On representational content and format in core numerical cognition." *Philosophical Psychology* 30(1-2), pp 119-139.

Barth, H. C., Kanwisher, N., and Spelke, E. (2003). "The construction of large number representation in adults." *Cognition* 86, pp 201- 221.

Barth, H. C., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., and Spelke, E. S. (2006). "Non-symbolic arithmetic in adults and young children." *Cognition* 98(3), pp 199-222.

Bates, E., Elman, J., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1998). *Rethinking innateness: A connectionist perspective on development.* Cambridge: MIT Press.

Beck, J. (2015). "Analogue Magnitude Representations: A Philosophical Introduction." *The British Journal for the Philosophy of Science* 66, pp 829-855.

Block, N. J. (1986). "Advertisement for a semantics for psychology." *Midwest Studies in Philosophy* 10(1), pp 615-678.

Brannon, E. M., and Terrace, H. S. (1998). "Ordering of the numerosities 1 to 9 by monkeys." *Science* 282, pp 746-749.

Burge, T. (2010). *Origins of Objectivity.* Oxford: Clarendon Press.

Burgess, N. (2006). "Spatial memory: how egocentric and allocentric combine." *Trends in Cognitive Science* 10(12), pp 551-557.

Calero, C., Shalom, D. E., Spelke, E. S., and Sigman, M. (2019). "Language, gesture, and judgment: Children's paths to abstract geometry." *Journal of Experimental Child Psychology* 177, pp 70-85.

Cantlon, J., and Brannon, E. M. (2006). "Shared system for ordering small and large numbers in monkeys and humans." *Psychological Science* 17(5), pp 401-406.

Carey, S. (2009). *The Origin of Concepts.* Oxford: Oxford University Press.

Carey, S. (2011). "Précis of The Origin of Concepts." *Behavioral and Brain Sciences* 34, pp 113-124.

Cheng, K. (1986). "A purely geometric module in the rat's spatial representation." *Cognition* 23, pp 149-178.

Cheng, K. (2008). "Whither geometry? Troubles of the geometric module." *Trends in Cognitive Sciences* 12(9), pp 355-361.

Cheng, K., Huttenlocher, J., and Newcombe, N. (2013). "25 years of research on the use of geometry in spatial reorientation: a current theoretical perspective." *Psychon Bull Rev* 20, pp 1033-1054.

Chrastil, E. R., and Warren, W. H. (2014). "From Cognitive Maps to Cognitive Graphs." *PLoS ONE* 9(11): e112544.

Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics.* New York: Oxford University Press. First edition published in 1997.

Dehaene, S. and Changeux, J. (1993). "Development of Elementary Numerical Abilities: A Neuronal Model." *Journal of Cognitive Neuroscience* 5, pp 390-407.

Dehaene, S. and Cohen, L. (1994). "Dissociable mechanisms of Subitizing and Counting: Neurophysiological Evidence from Simultanagnosic Patients." *Journal of Experimental Psychology: Human Perception and Performance* 20(5), pp 958-975.

Dehaene, S., Izard, V., Pica, P., and Spelke, E.S. (2006). "Core knowledge of geometry in an Amazonian indigene group." *Science* 311, pp 381-384.

Dillon, M., Huang, Y., and Spelke, E. S. (2013). "Core Foundations of Abstract Geometry." *Proceedings of the National Academy of Sciences* 110(35), pp 14191-14195.

Dillon, M., Persichetti, A., Spelke, E. S., and Dilks, D. D. (2017). "Places in the Brain: Bridging Layout and Object Geometry in Scene-Selective Cortex." *Cerebral Cortex* 2017, pp 1-10.

Dillon, M. and Spelke, E. S. (2015). "Core Geometry in Perspective." *Developmental Science* 18(6), pp 1-14.

Dillon, M. and Spelke, E. S. (2018). "From Map Reading to Geometric Intuitions." *Developmental Psychology* 54(7), pp 1304-1316.

Feigenson, L., and Carey, S. (2005). "On the limits of infants' quantification of small object arrays." *Cognition* 97(3), pp 295-313.

Flombaum, J. I., Junge, J., and Hauser, M. D. (2005). "Rhesus monkeys (*Macaca mulatta*) spontaneously compute large number addition operations." *Cognition* 97(3), pp 315-325.

Fodor, J. (1975). *The Language of Thought.* Cambridge: Harvard University Press.

Fuson, K. (1988). *Children's Counting and Concepts of Number.* New York: Springer.

Giaquinto, M. (2001a). "Knowing Numbers." *The Journal of Philosophy* 98(1), pp 5-18.

Giaquinto, M. (2001b). "What Cognitive Systems Underlie Arithmetical Abilities?" *Mind & Language* 16(1), pp 56-68.

Gray, E.R. et al. (2005). "Spatial encoding in mountain chickadees: features overshadow geometry." *Biol. Lett.* 1, pp 314-317.

Hale, B. and Wright, C. (2001). *The Reason's Proper Study.* Oxford: Oxford University Press.

Hafting, T., Fyhn, M., Molden, S., Moser, M., and Moser, E. (2005). "Microstructure of a spatial map in the entorhinal cortex." *Nature* 436(11), pp 801-806.

Hartley, T., Burgess, N., Lever, C., Cacucci, F., O'Keefe, J. (2000). "Modeling Place Fields in Terms of the Cortical Inputs to the Hippocampus." *Hippocampus* 10, pp 369-379.

Heck, R. (2000). "Cardinality, Counting, and Equinumerosity." *Notre Dame Journal of Formal Logic* 41(3), pp 187-209.

Hupbach, A., and Nadel, L. (2005). "Reorientation in a rhombic environment: No evidence for an encapsulated geometric module." *Cognitive Development* 20, pp 279-302.

Izard, V., Pica, P., Dehaene, S., Hinchey, D., and Spelke, E. S. (2011a). "Geometry as a Universal Mental Construction." In E. Brannon and S. Dehaene (Eds.), *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought, Attention and Performance.* Boston: Elsevier Academic Press, pp 319332.

Izard, V., Pica, P., Spelke, E. S., and Dehaene, S. (2011b). "Flexible intuitions of Euclidean geometry in an Amazonian indigene group." *Proceedings of the National Academy of Sciences* 108(24), pp 97829787.

Izard, V., and Spelke, E. S. (2009). "Development of sensitivity to geometry in visual forms." *Human Evolution* 23, pp 213-248.

Jenkins, C. S. (2005). "Knowledge of Arithmetic." *The British Journal for the Philosophy of Science* 56, pp 727-747.

Jenkins, C. S. (2008). *Grounding Concepts: An Empirical Basis for Arithmetical Knowledge.* Oxford: Oxford University Press.

Kanwisher, N., and Dilks, D. D. (2014). "The Functional Organization of the Ventral Visual Pathway in Humans." In L. Chalupa and J. Werner (Eds.), *The New Visual Neurosciences.* Cambridge: MIT Press.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science.* Cambridge: MIT Press.

Landau, B. (2017). "Update on 'What' and 'Where' in Spatial Language: A New Division of Labor for Spatial Terms." *Cognitive Science* 41 (Supp 2), pp 321-350.

Laurence, S. and Margolis, E. (1999). "Concepts and cognitive science." In E. Margolis and S. Laurence (Eds.), *Concepts: Core readings.* Cambridge: MIT Press, pp 3-82.

Laurence, S. and Margolis, E. (2002). "Radical concept nativism." *Cognition* 86, pp 25-55.

Lee, S. A., Sovrano, V. A., Spelke, E. S. (2012). "Navigation as a source of geometric knowledge: Young children's use of length, angle, distance, and direction in a reorientation task." *Cognition* 123, pp 144-161.

Lipton, J. S., and Spelke, E. S. (2003). "Origins of number sense: Large number discrimination in human infants." *Psychological Science* 15(5), pp 396-401.

Lipton, J. S., and Spelke, E. (2004). "Discrimination of large and small numerosities by human infants." *Infancy* 5(3), pp 271-290.

MacNamara, J. (1986). *Border dispute: The place of logic in psychology.* Cambridge, MA: MIT Press.

Maddy, P. (2007). *Second Philosophy: A Naturalist Method.* Oxford: Oxford University Press.

Maddy, P. (2014). "A Second Philosophy of Arithmetic." *The Review of Symbolic Logic* 7(2), pp 222-249.

Margolis, E. and Laurence, S. (2008). "How to learn the natural numbers: Inductive inference and the acquisition of number concepts." *Cognition* 106, pp 924-939.

McCrink, K., and Wynn, K. (2004). "Large-number addition and subtraction by 9-month old infants." *Psychological Science* 15(11), pp 776-781.

Moser, M., Rowland, D., Moser, E. (2015). "Place Cells, Grid Cells, and Memory." *Cold Spring Harb Perspect Biol* 7:a021808.

Nersessian, N. (1992). "How do scientists think? Capturing the dynamics of conceptual change in science." In R. Giere (Ed.), *Cognitive Models of Science.* Minneapolis: University of Minnesota Press, pp 3-44.

Newcombe, N., Ratliff, K., Shallcross, W., Twyman, A. (2009). "Young children's use of features to reorient is more than just associative: further evidence against a modular view of spatial processing." *Developmental Science* 13(1), pp 213-220.

O'Keefe, and Burgess, N. (1996). "Geometric determinants of the place fields of hippocampal neurons." *Nature* 381(6581), pp 425-428.

Parsons, C. (2009). *Mathematical Thought and Its Objects.* Cambridge: Cambridge University Press.

Pearce, J.M. et al. (2006). "Potentiation, overshadowing, and blocking of spatial learning based on the shape of the environment." *J. Exp. Psychol. Anim. Behav. Process.* 32, pp 201-214.

Piaget, J. (1936). *The Construction of Reality in the Child.* Abingdon: Routledge, 1954.

Prinz, J. J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis.* Cambridge: MIT Press.

Resnik, M. (1997). *Mathematics as a Science of Patterns.* Oxford: Clarendon Press.

Shea, N. (2011). "New concepts can be learned: Review of Susan Carey, *The Origins of Concepts.*" *Biology and Philosophy* 26, pp 129-139.

Shusterman, A., and Li, P. (2016). "Frames of reference in spatial language acquisition." *Cognitive Psychology* 88, pp 115-161.

Solstad, T., Boccara, C., Kropff, E., Moser, M., and Moser, E. (2008). "Representation of Geometric Borders in the Entorhinal Cortex." *Science* 322(19), pp 1865-1868.

Spelke, E. S. (2011). "Natural Number and Natural Geometry." In E. Brannon and S. Dehaene (Eds.), *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought, Attention and Performance.* Boston: Elsevier Academic Press, pp 287-317.

Spelke, E. S. and Kinzler, K. D. (2007). "Core Knowledge." *Developmental Science* 10(1), pp 89-96.

Spelke, E. S. and Lee, S. (2012). "Core Systems of Geometry in Animal Minds." *Philosophical Transactions of the Royal Society B* 367, pp 2784-2793.

Spelke, E. S., Lee, S. A., and Izard, V. (2010). "Beyond Core Knowledge: Natural Geometry." *Cognitive Science* 34, pp 863-884.

Spencer, J. P., Thomas, M., McClelland, J. (2009), *Toward a New Grand Theory of Development? Dynamic Systems Theory and Connectionism Reconsidered.* Oxford: Oxford University Press.

Starkey, P., and Cooper, R. (1980). "Perception of numbers by human infants." *Science* 210, pp 1033-1035.

Sutton, J. E., and Newcombe, N. S. (2014). "The hippocampus is not a geometric module: processing environment geometry during reorientation." *Frontiers in Human Neuroscience* 8(596), pp 1-6.

Thelen, E. and Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action.* Cambridge: MIT Press.

Twyman, A. and Newcombe, N. (2010). "Five Reasons to Doubt the Existence of a Geometric Module." *Cognitive Science* 34, pp 1315-1356.

Twyman, A., Newcombe, N., and Gould, T. (2013). "Malleability in the Development of Spatial Reorientation." *Developmental Psychology*, pp 243-255.

Winkler-Rhoades, N., Carey, S. C., and Spelke, E. S. (2013). "Two-year-old children interpret abstract, purely geometric maps." *Developmental Science* 16(3), pp 365-376.

Wood, J. N., Spelke, E. (2005). "Infants' enumeration of actions: Numerical discrimination and its signature limits." *Developmental Science* 8(2), pp 173-181.

Wynn, K. (1990). "Children's Understanding of Counting." *Cognition* 36, pp 155-93.

Xu, F., and Spelke, E. S. (2000). "Large number discrimination in 6-month old infants." *Cognition* 74, pp B1-B11.

Chapter 3

Visual Imagination and Mathematical Entitlement

Abstract: Recent work by Marcus Giaquinto on visual thinking in mathematics has provided us with a new appreciation for the role of visual imagination in the formation and warrant of basic beliefs in elementary arithmetic and geometry. Yet, Giaquinto remains skeptical about the epistemic role that visual imagination can play in elementary analysis. He claims that our visual faculties are unreliable guides to the "limiting behavior" of certain continuous functions. But a recent branch of mathematical logic, namely o-minimality, takes as one of its primary objects *definable continuous functions*. I will argue that our visual faculties are dependable guides to the behavior of these objects. More specifically, I will argue that the o-minimal Intermediate Value Theorem admits of imagery-based mathematical entitlement despite Giaquinto's (sound) argument that the real-analytic IVT cannot be known in this way.

3.1 Introduction

In his 2007 book, *Visual Thinking in Mathematics*, Marcus Giaquinto provides an answer to a question raised by Plato: "how is pure geometrical knowledge possible?" (p 12). Giaquinto argues that a proper understanding of visual imagination and geometrical concepts can explain the formation of true warranted geometrical beliefs. Can a parallel account be given for mathematical knowledge outside of elementary geometry? Giaquinto maintains that theorems in real analysis and other advanced mathematical disciplines cannot be known through visual thinking.[1] In this essay I will argue that Giaquinto underestimates the potential explanatory power of his own epistemic framework. By focusing on the case of the Intermediate Value Theorem, I will show that visual thinking can generate mathematical knowledge in advanced disciplines like model theory and topology.

The primary reason Giaquinto develops his account of visual thinking is to refute a long-standing position (held by some prominent mathematicians and philosophers) that visual imagination and diagrammatic reasoning are too unreliable to produce warranted mathematical beliefs. In Giaquinto's own words:

> A time-honoured view, still prevalent, is that the utility of visual thinking in mathematics is only psychological, not epistemological. ...Visual representations...cannot be a resource for discovery, justification, proof, or any other way of adding epistemic value to our mathematical capital – or so it is held. The chief aim of this book is to put that view to the test. I will try to show how, why, and to what extent it is mistaken. (2007, p 1)

Giaquinto takes aim here at numerous claims by both mathematicians and philosophers over the past 200 years, such as Russell's assertion that arithmetic and geometric "self-evidence is often a mere will-o'-the-wisp, which is sure to lead us astray" (1901, p 78). Arguably, this skeptical tradition can be traced to 1817, when Bernardo Bolzano appealed to the real-analytic Intermediate Value Theorem (IVT) to defend his view that our visual faculties are not reliable guides to the behavior of ϵ-δ continuous functions. In the early

[1] When the representational contents of visual imagination give rise to mathematical beliefs in a chain of reasoning, Giaquinto says that this reasoning constitutes "visual thinking" (2007, p 1).

20th century, Russell (1919) and Hahn (1933) put forward an historical narrative according to which 19th-century mathematics is a story of a "flight from intuition." They claim that these mathematicians continually uncovered cases where visual imagination leads us to believe various mathematical falsehoods.[2] According to Russell and Hahn, these demonstrations eventually convinced the mathematical community that a search for the foundations of mathematical knowledge was necessary. The first major victory of this movement was the discovery of a "rigorous" definition of the notion of a *continuous function.* According to this narrative, once ϵ-δ definition of continuity was formulated, true mathematical knowledge of the IVT became possible. It is this historical narrative that makes the IVT particularly worthy of study for anyone concerned with visual thinking in mathematics, and it is why I focus on the IVT in this essay.

In Part 2, I distinguish some basic epistemological notions. In Part 3, I describe a plausible set of geometric contents for visual imagination according to the current dominant theory of working memory. In Part 4, I consider a single proposition of elementary geometry – *for any square, the parts either side of its diagonal are congruent* – and show how we can acquire knowledge of this proposition through visual thinking. In Part 5, I will extend this account of mathematical knowledge to the o-minimal Intermediate Value Theorem. Unlike the real-analytic Intermediate Value Theorem, which Giaquinto has discussed extensively (1994, 2007, 2011), I will argue that the o-minimal version can be known through visual thinking. In Part 6, I step back from the IVT and show that the developers and practitioners of o-minimality were engaged in a general epistemological project. These mathematicians intended the primary objects of o-minimality to be inherently visualizable. I will consider the extent to which they succeeded by asking if our visual faculties are reliable guides to the behavior of these objects.

[2]One can find discussions of the falsehoods that allegedly powered the flight from intuition in the following sources: Russell (1901), p 94; von Koch (1906); Forder (1927), p 46; Coffa (1982), p 686. Cf. Norman (2006), Mancosu (2005). It should be noted, however, that the 'flight from intuition' narrative has been refuted by historians of mathematics (e.g., Gray 2008, Heis 2008, Lützen 2003). Nevertheless, it continues to hold weight among some philosophers.

3.2 Imagery-based Mathematical Entitlement

Marcus Giaquinto's mathematical epistemology focuses on a very specific type of knowledge that requires what I will call *imagery-based mathematical entitlement*. In this section I will flesh out what I mean by this phrase and the more basic notion of *epistemology*. Epistemology, in the broadest sense, concerns a person's *right* to accept a proposition or theory as true.[3] Some rights to accept a proposition p are partially grounded in the believer's ability to produce a proposition q as a *reason* (i.e., justification) for the belief. Other rights to accept a proposition are not grounded in such an ability, and such rights are called *entitlements*.[4] When a justified or entitled belief is true, I will refer to it as *knowledge*. The most familiar type of entitlement is a *perceptual entitlement*. A perceptual entitlement is a right to believe a proposition that is grounded in (a) the reliability of the believer's perceptual systems in generating perceptual states with accurate contents, and (b) a deductively valid connection between the contents of the underlying perceptual states (e.g., *this triangle is scarlet*) and the contents of the belief that the perceptual states generate (e.g., *this figure is red*).

Giaquinto maintains that for his target epistemological case studies, namely mathematical beliefs generated from visual imagination (e.g., *the angles of a triangle sum to 180°*), our right to believe should not theorized in terms of perceptual entitlement.[5] Mathematical beliefs, unlike perceptual beliefs, are not about the believer's visible environment and so possession of a systematically accurate perceptual system does not add any epistemic force to the belief. In other words, condition (a) above is not necessary for the mathematical

[3] According to this understanding, knowledge and justification are merely subtopics within epistemology.

[4] The terminology of entitlements comes from Burge (1993, 1998, 2003), Dretske (2000), and Goldman (2001).

[5] Giaquinto uses the word "discovery" instead of "entitlement," which I assume he draws from Lakatos (1976). The only difference is that *discovery* excludes entitlements from testimony and entitlements that involve a "violation of epistemic rationality" (2007, p 2). I will be exclusively concerned with mathematical entitlement in this essay, so the reader should assume these restrictions hold unless otherwise stated.

kind of entitlement that Giaquinto theorizes. For this reason, imagery-based mathematical entitlements (as distinguished from perceptual entitlements) are *apriori* entitlements.[6]

That is, these beliefs do not derive their epistemic status from a reliably-produced perceptual state. Additionally, since the mathematical belief is generated from visual imagination rather than perception, condition (b) is not necessary either. However, a validity condition similar to (b) is necessary for imagery-based mathematical entitlement: the ground of one's right to believe that, e.g., *for any square the parts either side of its diagonal are congruent*, is a deductively valid connection between the underlying imagery contents and the generated belief contents. In Part 4, I will flesh out the imagery contents from which this proposition about squares may be deduced. Before diving into this particular example, I will use Part 3 to describe what we know about the geometric contents of visual imagery. current cognitive-scientific understanding of mental imagery and its contents. For now, all one needs to know is that when these geometric contents produce a mathematical belief, we will say that the resulting belief is *imagery-based.*

The third and final kind of entitlement that I will mention here is *analytic entitlement.* This kind of entitlement accrues to mathematical beliefs such as *two is a number* and *every square has four sides*, as well as non-mathematical beliefs such as *every apple is a fruit* and *my house lights are either on or off.*[7] In every such case, anyone who possesses the component concepts of the proposition will necessarily find the proposition self-evident.[8] In

[6]Giaquinto (2007), pp 2, 8. Cf. Burge (1993), p 274. For discussion of non-imagery-based kinds of apriori entitled beliefs (e.g., introspective, analytic, or logical beliefs), see the essays cited in footnote 4.

[7]For endorsements of the idea that our entitlement to certain set-theoretic axioms is analytic, see Kreisel (1967, p 144), Gödel (1964, p 260ff), and Fraenkel (1927, p 61).

[8]I would like to note, somewhat tangentially, a minor difference between Giaquinto's theory of analytic entitlement and my own. Giaquinto elaborates the idea of analytic entitlement in terms of Christopher Peacocke's (1992) theory of concept possession, which is a theory to which I do not want to commit myself. According to Peacocke's theory, possession of a concept is nothing more than acceptance of a set of "primitively compelling" propositions and inference rules. I do not wish to commit myself to this theory of concept possession, but I find a weaker claim plausible: for some concepts, acceptance of certain propositions is necessary (but not sufficient) for possession of the concept. For arithmetic and geometric concepts, this claim seems exceptionally plausible. My rejection of Peacocke's theory and acceptance of the weaker claim aligns with Prinz (2002), who claims that Peacocke's theory "seems to work well for certain taxonomic and kinship terms,...mathematical and theoretical terms,...and artifact concepts" (p 32). Prinz goes on to argue

contrast to imagery-based mathematical entitlement, one's right to believe these propositions is not grounded in any facts about pre-belief representational systems (e.g., the existence of a valid connection between visual imagery and the beliefs it generates). Analytic entitlement will only become significant in this essay when I elaborate the notion of a "valid connection," which is necessary for imagery-based mathematical entitlement in Part 4. In particular, when we establish one example of a deductively valid connection between the contents of visual imagery and the resulting belief, we will allow premises that are self-evident propositions *in addition to* the premises that are imagery contents.

Before I move on to characterizing the geometric contents of visual imagination, I will briefly specify what 'mathematical' means in *imagery-based mathematical entitlement*. My understanding of mathematics is broad. My use of this term not only rejects logicist claims such as, "formal logic...is the same thing as mathematics" (Russell 1901, p 74), but it also rejects more lenient uses of 'mathematics' where *professional* mathematicians are the dominant concern (Paseau 2014). In contrast, my notion of mathematics is just as much about the representational content and inferential activity associated with K-12 mathematics textbooks and exercises as it is about the proofs in mathematics journals.[9] Giaquinto argues that elementary beliefs in number theory, algebra, and geometry admit of imagery-based mathematical entitlement. But Giaquinto also argues that this kind of entitlement cannot accrue to beliefs in analysis. Giaquinto's argument leaves the reader with the impression that imagery-based entitlement is impossible for mathematical beliefs about more 'advanced' mathematical subject matters, such as algebraic topology, set theory, complex analysis, geometric model theory, analytic number theory, and algebraic geometry.[10]

that Peacocke's account is not be adequate for every kind of categorization capacity that we might want to call conceptual, such as dot-pattern categorization (p 51), color categorization (p 52), bird categorization (p 53-55). A prototype theory of concept individuation (Rosch and Mervis 1975) may be preferable in these cases. It is possible that Giaquinto also believes this since he appeals to Kosslyn's notion of a *category pattern*, which is a prototype (Kosslyn 1994, p 180).

[9]This seems to be how Giaquinto (2005b) understands the delineation of mathematics as well.

[10]Indeed, most philosophers of mathematics place similar epistemic limits on visual imagination and diagrams in 'advanced' disciplines. One exception is Brown's (2008) argument that the analytic Intermediate Value Theorem can be diagrammatically known. However, I believe Giaquinto (2011) provides a convincing

In Part 5, I will show that this impression is misleading. Before I address the question of which mathematical subdisciplines admit imagery-based mathematical entitlement, I will first need to describe the geometric contents of visual imagination that make imagery-based mathematical entitlement possible.

3.3 The Geometric Contents of Visual Imagination

To articulate his view of the contents of visual imagery, Giaquinto appeals to the work of cognitive psychologist Stephen Kosslyn. Kosslyn's theory of visual imagery remains more or less the dominant paradigm in cognitive psychology today.[11] According to Kosslyn's theory, when a human generates a mental image of, e.g., a square, a tent, or a knife, their visual imagination produces three kinds of representations: an environment-centered coordinate system, a set of object-centered intrinsic axes, and a set of axis-dependent spatial relations. I will describe each kind of representation in the following paragraphs.[12]

According to Kosslyn, when we imagine an object we attribute a coordinate system to the space surrounding the object.[13] Perhaps the most widely accepted argument for this claim comes from the experimental results of Roger Shepard. Roger Shepard's studies of mental rotation (Shepard and Metzler 1971, Cooper and Shepard 1973) are widely regarded

refutation of his view. Another exception, one which I find more promising, is De Toffoli and Giardino (2014).

[11]Endorsements of Kosslyn's theory of imagery contents can be found in Bettencourt and Xu (2016), Christophel et al. (2015), Pearson et al. (2013), Hamamé et al. (2012), and Logie and van der Meulen (2009). In cognitive psychology and cognitive neuroscience today, discussion of visual imagination is often couched within a larger discussion of working memory. For this reason, Baddeley's *episodic buffer* is more frequently cited than Kosslyn's *visual buffer*. I maintain, along with Pearson (2001) and Logie and van der Meulen (2009), that these two terms refer to the same buffer.

[12]For discussion of and evidence for this coordinate structure, see Kosslyn (1994, ch 7). All of my examples of an imagined object – a square, a tent, and a knife – are taken from Kosslyn (1994).

[13]The origin and axes of this coordinate system (relative to the visualizer) are determined by factors such as "the gravitational upright,...position on the retina, and the way the body is oriented" (Kosslyn 1994, p 132).

as revealing the existence of an environment-centered coordinate system in mental imagery.[14] In the first study, Shepard and Metzler asked participants whether two distinct 3D tetris figures at various orientations have the same shape. They found that the decision times of participants increased linearly with the angular distance of the necessary rotation. This result has been confirmed many times, and parallel results have been found for other spatial configurations such as letters (Cooper and Shepard 1973) and 2D line figures (Hochberg and Gellman 1977, Jordan *et al.* 2001).[15] Kosslyn uses Shepard's result to claim that we attribute an environment-centered coordinate system to the space around an imagined object. If this were not true,

> mental rotation would not be necessary. In fact, one of the reasons Shepard and Metzler's (1971) original mental rotation experiment had such a large impact was that the then-current approaches to visual representation emphasized object-centered representations, and the rotation results were clearly not predicted by such approaches. (Kosslyn 1994, p 149)

Kosslyn here rejects an object-centered approach to mental rotation, which holds that we merely attribute *intrinsic* axes to an object. These intrinsic axes are primarily determined by the symmetry properties of an object. So, for example, the intrinsic axes of a square will be the lines that segment the square into four congruent square parts. Without an additional set of *environmental* axes, the object-centered approach is unable to give an account of what

[14]For instance, Jeffrey Zacks begins his 2007 meta-analysis of mental rotation studies with these assertions: "Mental rotation is a hypothetical psychological operation in which a mental image is rotated around some axis in three-dimensional space. Mental rotation was first revealed in behavioral experiments (Cooper and Shepard 1973, Shepard and Metzler 1971) in virtue of a striking finding: The time to make a judgment about a rotated object often increases in a near-linear fashion with the amount of rotation required to bring the object into alignment with a comparison object or with a previously learned template." (Zacks 2007, p 1)

[15]Additionally, fMRI studies have found that mental rotation tasks are accomplished by spatiotopic regions of the posterior parietal cortex and the activation levels in these regions increases linearly with angular distance (Harris and Miniussi 2003, Zacks 2007). Together, these behavioral and neuroimaging results "provide relatively good support for the hypothesis that mental rotation is a continuous transformation" of an imagined object within a stable, environment-centered coordinate system (Zacks 2007, p 6).

it would mean to *continuously* transform a square by mentally rotating it 45°.[16] Hence, Shepard's result is often taken to suggest that we represent both an environment-centered coordinate system and a set of intrinsic axes, the latter of which I will call a *reference frame*.

Beside Shepard's result, the existence of an object-centered reference frame (in addition to an environment-centered coordinate system) in visual imagination is further suggested by the symmetry experiments such as Palmer and Hemenway (1978) and McMullen and Farah (1991).[17] Palmer and Hemenway found that decision times for whether a figure has a reflection symmetry (e.g., symbols 'H', '%') across an intrinsic axis is substantially shorter than the decision time for whether a figure has a rotational symmetry (e.g., the letter 'Z'). Their explanation for this result is that reflection symmetries, but not rotational symmetries, are encoded into the very representation of the figure. Additionally, McMullen and Farah found "that the effects of orientation on naming time are eliminated with practice for symmetrical, but not asymmetrical, pictures" (Kosslyn 1994, p 133). They also explained their result by claiming that reflection symmetries are explicitly encoded in the representation of imagined and perceived objects.[18]

The third kind of representation that Kosslyn posits in visual imagination is a set of axis-dependent spatial relations. The reflection symmetries mentioned above are just one example of this kind of representation. In addition, Kosslyn posits representations of metric relations (e.g., distance, parallelity) and qualitative relations (e.g., collinearity, co-termination, 'on', 'left of', 'connected to', 'above') within visual imagination.[19] The

[16]An account of mental 'rotation' in terms of discrete changes, such as using intrinsic axes that segment the square into four triangles rather than four squares, is ruled out by Shepard's result.

[17]Kosslyn's (1994, pp 133-136, 160) own endorsement of intrinsic axes draws on these experiments and the related symmetry experiments of Corballis and Beale (1976, 1983), Bower and Glass (1976), Palmer (1977), Jolicoeur and Kosslyn (1983), McMullen and Farah (1991).

[18]I believe symmetry results like these were a major reason that Kosslyn abandoned his early theory of imagery contents (Kosslyn and Shwartz 1977, p 278), according to which shapes representations are encoded by a two-dimensional grid of filled and unfilled cells. The symmetry theory does not require such a rich grid structure to be represented (though, of course, even the symmetry theory acknowledges that the non-representational retinal image will have a 'grid structure' in a sense). In philosophy, Michael Tye (1991, chs 5, 6) has incorporated Kosslyn's older grid structure into his theory of imagery contents.

[19]Kosslyn (1994), pp 108-109, 227-228.

experimental support for the existence of these spatial relations has been even less conclusive than the available evidence discussed above for the coordinate system and reference frame. For this reason, I will not review these results here. Nevertheless, I find Kosslyn's theory of imagery content plausible enough given the currently available behavioral and neural evidence. Hence I will follow Giaquinto (2005a, 2007) in accepting these contents for the rest of this essay while also noting that (i) this set of representations "is not the only one that is consistent with the data" (2007, p 28), and (ii) our account of imagery-based mathematical entitlement will not be significantly altered if, say, *collinearity* is not represented in visual imagination.

Finally, I should note that in addition to the three kinds of representations above, Kosslyn (1994, ch 10) proposes the existence of a fixed set of transformations – rotation, translation, similarity (i.e., 'zooming'), reflection, and stretching – that a visualizer has at their disposal to voluntarily manipulate the representational contents of visual imagination. Kosslyn describes some additional decision-time experiments in favor of the existence of these mental transformations, but caveats (i) and (ii) also apply to the existence of these transformations. Given the above notion of visual imagination, we will now turn our attention to Giaquinto's illustrative example of a geometric belief with imagery-based mathematical entitlement.

3.4 Geometric Knowledge through Visual Imagination

With the above epistemological definitions and imagery contents fixed, I will focus on one example of imagery-based mathematical entitlement from Giaquinto (2007) that will serve as a guide for later cases. Consider the following claim:

(S) For any square x and for any diagonal k of x, the parts of x either side of k are congruent. (Giaquinto 2007, p 38)

After visualizing a square with a diagonal through it, a person may come to believe (S) without being able to give a justification for the belief.[20] Nevertheless, the person may be *entitled* to believe (S) if the belief was acquired through a valid connection with visual imagination. What constitutes this valid connection? First, their concept *square* must have the correct "category specification." That is, a certain set of imagery contents must cause the application of the concept *square* to the imagined object. Giaquinto proposes a category specification for square that is empirically plausible:

> Category Specification for *square*
>
> Let V and H be the vertical and horizontal axes of the reference frame. Then, to categorize a figure as a square, it is necessary and sufficient that visual imagination represents the following features:
>
> - Plane surface region, enclosed by straight edges: edges parallel to H, one above and one below; edges parallel to V, one each side.
> - Symmetrical about V.
> - Symmetrical about H.
> - Symmetrical about each axis bisecting angles of V and H.[21]

This category specification must not only guide the application of the concept *square*. It must also generate a visual image with these contents when one visualizes a square. A person

[20]Besides its intuitive support, the idea that working memory (and hence visual imagination) triggers the formation of beliefs has some experimental support in the case of the phonological loop (Baddeley *et al.* 1987, Baddeley *et al.* 2009).

[21]Giaquinto (2007), p 23. He draws these contents primarily from Stephen Palmer's experiments on perceptual judgments of shape, orientation, and symmetry (Palmer and Hemenway 1978, Palmer 1980, 1983, 1985). For Giaquinto, these contents are also supposed to arise in perception. But, as argued in Chapter 1, I do not think that perceptual content includes a coordinate system. Nevertheless, it seems plausible that these contents may be stored in a post-perceptual memory system, such as Logie's visual cache (Logie 1995, Logie and Cowan 2015). But as Palmer (1999) notes, "the nature of shape perception is so complex and enigmatic that there is as yet no accepted theory of what shape is or how shape perception occurs" (p 327).

with imagery-based entitlement to (S) must have the correct category specification for the concept *square*. However, they must *also* have the correct category specifications for all other visual concepts in (S), including *diagonal*, *part*, and *congruence*.[22]

The second requirement for imagery-based entitlement to (S) is ability to reflect the square over its imagined diagonal. This ability is constitutive of a properly functioning imagery system. Reflecting the square over its diagonal should then cause the concept *congruence* to be applied to the parts of the square either side of the diagonal. The final requirement for imagery-based entitlement to (S) is the disposition to generalize over all imaginable square figures through imagery transformations.[23] us to believe exercise a disposition a If rotation, translation, and similarity are constitutive of visual imagination as Kosslyn claims, then this requirement is also met by anyone with a properly functioning imagery system.[24]

The above three requirements are sufficient for imagery-based entitlement to (S) because they establish a deductively valid connection between the imagery contents and the generated belief (S). This valid connection can be written in argument form as follows:

<u>Valid Connection for (S)</u>

[22]More precisely, these concepts are *diagonal of a quadrilateral*, *part of a figure*, and *congruence of two figures*.

[23]Giaquinto (2007, pp 40-42) maintains that this disposition to generalize over all squares is the very same disposition that can causes us to believe other universal mathematical claims about other shapes and configurations. If this psychological hypothesis turns out to be correct, Giaquinto is right to assert that this disposition will only generate knowledge if it is *reliable* (i.e., consistently generates true beliefs except for cases of malfunction).

[24]It is possible that this generalization over the set of all imaginable square figures is not the result of mental transformations, but simply the results of recognizing that all square-with-diagonal configurations "share all their intrinsic geometric properties" (Giaquinto 2007, pp 137, 154). Other types of generalization that Giaquinto posits do explicitly draw on Kosslyn's set of mental transformations. For instance, Giaquinto claims that we generalize over all *triangular* configurations of dots irrespective of the number of dots in each row. Giaquinto claims that the *phenomenal experience* of a dot pattern can be numerically indeterminate and thus we are disposed to make a shape-based generalizations based on similarity and rotational transformations (2007, pp 144, 156). This strikes me as *prima facie* plausible, but I will largely remain silent on arithmetical knowledge in this essay. Moreover, it is not necessary for us to determine here which of these two accounts of geometric generalization. Either way, the three requirements I have placed on imagery-based mathematical entitlement will be very similar.

P1. Every square is symmetric about its diagonals. (Specification for *square*)

P2. Every quadrilateral that is symmetric about its diagonal has two parts on either side of the diagonal that precisely overlap after a mental reflection. (Structure of visual imagination)

P3. Every pair of figures that can be made to precisely overlap with another through translation, reflection, and rotation are congruent. (Specification for *congruent*)

C. Therefore, for any square with a diagonal, the parts either side of its diagonal are congruent.

This deductively valid argument establishes a valid connection between (S) and visual imagination. It is important to note that this argument is not psychologically implemented by the person with imagery-based entitlement to (S). In fact, it is conceivable that someone entitled to (S) cannot even *entertain* any of the three premises in this argument. For instance, they may not possess concepts like *symmetry* or *reflection* that are necessary to entertain these premises. But even if they are able to think these premises and believe them, psychologically implementing this argument would not produce imagery-based entitlement to (S). At best, it would produce a justification to believe (S). Unlike this argument, the psychological route to an imagery-generated belief in (S) that I describe above involves: generating a particular imagined object, manipulating that imagined object, applying the concept *congruence* to the parts of that imagined object, and generalizing over all imaginable squares.[25] The argument above is not a description of psychological reality. Nevertheless, it establishes a valid connection between visual imagination and the generated belief. According to my theory of entitlement, this is sufficient for imagery-based mathematical entitlement to (S).

[25]This sequence of psychological acts reveals a general fact about the mathematical *productivity* of visual imagination. The heart of this productivity "lies in viewing a form in two ways at once" (Giaquinto 2007, p 158). In this case, we view a figure as both a square and as a composition of two triangles. Giaquinto also notes that cognitive scientists have theorized seeing in "two ways" (e.g., a rotated 'D' on top of a 'J' will be categorized as an umbrella) and often they explain these phenomena in terms of Kosslyn's visual buffer (Finke *et al.* 1989, p 76).

Which mathematical subject matters are most vulnerable to imagery-based mathematical entitlement? Clearly, claims about rectilinear figures are good candidates for this type of knowledge. But other mathematical concepts have category specifications as well. Giaquinto (2007) makes a number of compelling arguments that integers, algebraic structures, curves, and functions all have category specifications. This makes mathematical beliefs about these objects potentially subject to imagery-based entitlement. Below I will focus on planar curves and continuous functions. I will argue that certain concepts about these objects have category specifications. I will use the above square-and-diagonal example as a guide for demonstrating that certain beliefs about these curves and functions admit of imagery-based mathematical entitlement. In particular, my focus will be on the Intermediate Value Theorem. After I have detailed Giaquinto's argument for an *invalid* connection between visual imagination and the real-analytic IVT, I will consider whether the parallel argument holds in the case of the o-minimal IVT.

3.5 Knowing the o-minimal IVT through Visual Imagination

Bolzano and many mathematicians after him, used the Intermediate Value Theorem to cast doubt on the reliability of using diagrams to form mathematical beliefs. In other words, Bolzano questions the *epistemic status* of imagery-generated belief in the IVT. He argues that using a diagram in this case gives rise to a false, intermediate belief. Is Bolzano right? It depends on how we formulate the content of the IVT. I will try to answer this question for three different formulations of IVT: the synthetic formulation, the real-analytic formulation, and the o-minimal formulation. I will show that, in addition to the synthetic IVT, the o-minimal IVT admits of imagery-based mathematical entitlement.

3.5.1 The Synthetic IVT

The simplest and most intuitive formulation of the Intermediate Value Theorem is called the *synthetic* IVT. This formulation is a universal statement over the concept *planar curve*. The category specification for *planar curve* is as follows:

> Let y and x be the vertical and horizontal axes of the reference frame. Then, to categorize a figure as a planar curve, it is necessary that visual imagination represents the following features:
>
> - Bounded contour with no visible gaps[26]
> - Piecewise monotonic (i.e. finitely oscillating)[27]
> - Piecewise smooth (i.e. has a tangent at all but finitely many points)[28]
> - Passes the vertical line test (i.e., no point is directly above another)

Under this category specification, a planar curve will never have a segment that cannot be visualized. In other words, a planar curve will never exhibit such "pathological behavior" as filling a two-dimensional space (e.g., the "Hilbert curve"), having an infinitely-oscillating segment (e.g., the "topologist's sine curve"), or having infinitely many singular points (e.g., the "Weierstrass function").[29] With this in mind, we can state the synthetic IVT as follows:

> (SI) Any planar curve that goes from below the x-axis to above it, intersects the x-axis. (Giaquinto 2011, p 298)

Giaquinto maintains that, if a person possesses the concept *planar curve* and has the correct category specification of it, visual imagery may disposed them to believe (SI). He does

[26]Giaquinto (2011), pp 297-298. Giaquinto also refers to the property of having no visible gaps as "graphical continuity."

[27]Giaquinto (2007), p 185, fn 12.

[28]Giaquinto (2011), p 299. Additionally, Giaquinto speaks of "the 'intuitive' belief that a continuous function must have a derivative everywhere except at isolated points" (Giaquinto 2007, p 3).

[29]I intended "planar curve" to be synonymous with Giaquinto's (2007, 2011) term "pencil-continuous curve."

not discuss the psychological mechanism by which this generalization over all planar curves would normally occurs. For our purposes, we may assume that the *stretching* and *similarity* transformation of visual imagination is responsible for generating belief like (SI). More importantly, this imagery-generated belief has a valid connection to visual imagination. Hence, it has imagery-based mathematical entitlement. We can write this valid connection in argument form as follows:

<u>Valid Connection for (SI)</u>

P1. All planar curves are piecewise smooth and monotonic. (Specification for *planar curve*)

P2. All piecewise smooth and monotonic contours from below the x-axis to above have a segment that is smooth, monotonic and goes from below the x-axis to above. (Structure of visual imagination)

P3. All smooth, monotonic curve segments that go from below the x-axis to above it will *intersect* the x-axis at a point. (Structure of visual imagination)

C. Therefore, any planar curve that goes from below the x-axis to above it, intersects the x-axis.

Similar to the case of (S) above, this valid connection between (SI) and visual imagination is the ground of one's imagery-based entitlement to believe (SI). Can additional valid connections be established for other formulations of the Intermediate Value Theorem? Giaquinto makes a convincing argument that no valid connection exists between visual imagination and the analytic IVT. I want to review Giaquinto's argument before attempting to answer this question for the o-minimal formulation of the IVT.

3.5.2 The Analytic IVT

When Russell, Hahn, and Bolzano rejected visual imagination as a source of knowledge in mathematics, they all cite the analytic formulation of the IVT as a major motivation.

Although Giaquinto disagrees with their conclusion, he does admit that the analytic IVT cannot be known through visual thinking. I agree. The sheer symbolic complexity of the analytic IVT, as compared to the synthetic IVT, suggests by itself that this theorem may be about objects that cannot be visualized:

> (AI) If f is an ϵ-δ continuous function on $[a, b] \subset \mathbb{R}$ and $f(a) < 0 < f(b)$, then $f(c) = 0$ for some $c \in (a, b)$.[30]

Let us assume for now that visual imagination can generate a belief in (AI). If so, the concept ϵ-δ *continuous function* must have a category specification. Plausibly, it will be the very similar to the category specification of the concept *planar curve*. But this already poses a problem for the analytic IVT case. As we saw above, planar curves do not exhibit certain pathological behaviors like oscillating infinitely many times or having an infinite number of singular points. Yet, some notable ϵ-δ continuous functions demonstrably do exhibit these pathological behaviors. At this point, we may try to revise the category specification of ϵ-δ *continuous function* to include these functions. However, these pathological functions are notable precisely because they cannot be specified in visual imagination. Hence, there is no *correct* category specification for the concept ϵ-δ *continuous function*.[31] This fact makes it impossible to form universal beliefs about ϵ-δ continuous functions with imagery-based entitlement. If we were to attempt to write out a valid connection for (AI) as we did for (SI) above, we would fail. The parallel first premise – "all ϵ-δ continuous functions are

[30]Cf. Rudin (1976), p 93; Giaquinto (2011), p 297, fn 28. This is, in fact, the Intermediate Zero Theorem. For purposes of argumentation, I will assume, along with Giaquinto (1994, 2007, 2011) and Brown (2008), that we can use the IZT without loss of generality. Also, note that the range of this function is the metric space $\mathbb{R}$.

[31]To be clear, I am not denying that we might associate a set of *prototypical* visual features with this concept. In fact, Kosslyn (1994, p 178) calls this set of prototypical features an *exemplar pattern*. Exemplar patterns may give rise to universal beliefs about the associated concept, but these beliefs will lack a deductively valid connection to visual imagination. In other words, such beliefs will lack imagery-based entitlement. Kosslyn gives the example of an exemplar pattern of the concept *dog*. This pattern is a set of visible features of a particular dog, Rover. When one imagines Rover, one may disposed to form various universal beliefs about dogs. But these beliefs would not have imagery-based entitlement because Rover's features are not had by all dogs. There is no valid connection. Of course, such universal beliefs may have inductive justifications, but this is another matter entirely.

piecewise smooth and monotonic" – is nowhere to be found in the category specification for ϵ-δ *continuous function*. Without this premise, we cannot write a valid connection between the analytic IVT and visual imagination. Giaquinto expresses this same point when he claims that the analytic IVT does not admit of imagery-based entitlement:

> The problem is that there are some functions ϵ-δ continuous on a closed interval ... [that] do not have curves. Any continuous nowhere-differentiable function is an example. So generalizing from the diagram here is not valid. (2011, p 300)

The analytic IVT is not the only mathematical theorem that eludes imagery-based entitlement.[32] Indeed, it may be tempting to think that for all but the most elementary mathematical theorems, imagery-based entitlement will be untenable. Giaquinto himself expresses this view when he claims that the reason (AI) does not admit of imagery-based entitlement is that it "depends on theoretical concepts from e.g. real (or complex) analysis, or something similarly abstruse" (2007, p 29). But I see no reason why complex and 'abstruse' mathematical concepts and theorems should be excluded from imagery-based entitlement. In the rest of this paper, I want to push back against this tidy epistemological partition of mathematics by showing that the 'abstruse' o-minimal formulation of the IVT has a valid connection to visual imagination.

[32] Giaquinto gives two other examples of mathematical beliefs that do not admit of imagery-based entitlement. One example comes from Kenneth Manders' discussion of the distinction between exact and co-exact properties in Euclidean practice (Manders 1996, 2008a, 2008b). In this setting, a diagram of a triangle may dispose one to believe a false universal statement because the diagram contains a co-exact property that "depends on" exact properties (Giaquinto 2011, pp 291-293). The second example is a generalization over paths in Cantor space (Giaquinto 2008b; 2007, ch 11).

Although Giaquinto does not discuss it, there is a fourth formulation of the Intermediate Value Theorem. This *topological* IVT, formulated and proved by Munkres (2000, p 154) as Theorem 24.3, is not subject to imagery-based entitlement. As was the case for ϵ-δ continuous functions, some topologically continuous functions cannot be visualized. A complete study of epistemology of the IVT should include considerations of all four of these theorems (and probably more). Recognizing the variety of IVTs is crucial for dissolving the presumption that the IVT belongs exclusively to real analysis.

3.5.3 The O-minimal IVT

The o-minimal formulation of the IVT is able to overcome the two difficulties that plagued the analytic IVT in the last section. First, the concept *o-minimal continuous function* is capable of having a category specification, unlike the concept ϵ-δ *continuous function*. Second, there is a valid connection between the o-minimal IVT and visual imagination. Unlike the analytic IVT, we can write this valid connection in a deductive argument in a way that parallels the argument for (SI) above. This section is aimed at establishing both of these claims. We will begin by stating what exactly o-minimality is before returning to these claims.

O-minimality is a branch of model theory. I will review some general model-theoretic definitions introducing the concept *o-minimal continuous function*. This will help us to grasp the content of the o-minimal formulation of the IVT. At the most basic level, model theory begins with a collection of formal languages:

Definition 3.5.1. A model-theoretic **language** $\mathcal{L}$ is given by specifying the following data:

a) A set of function symbols $\mathcal{F}$ and non-negative integers n_f for each $f \in \mathcal{F}$.

b) A set of relation symbols $\mathcal{S}$ and positive integers n_S for each $S \in \mathcal{S}$.

c) A set of constant symbols $\mathcal{C}$.[33]

For example, the language of groups is specified by two symbols: the identity element 'e' and the group operation '$\cdot$'. Other language may contain, e.g., a function symbol for every analytic function on $\mathbb{R}$. A model theorist might consider one language in isolation or, as in the case of o-minimality, many languages all at once. For each language, there is a collection of structures on which the symbols of the language are given a meaning. More precisely:

[33] Marker (2002b), p 8.

Definition 3.5.2. A model-theoretic **$\mathcal{L}$-structure** $\mathcal{R} = (R, (S_i^{\mathcal{R}})_{i \in I}), (f_j^{\mathcal{R}})_{j \in J})$ consists of a nonempty set R, the relations $S_i^{\mathcal{R}}$ each of which is the interpretation of $S_i \in \mathcal{S}$, and the functions $f_j^{\mathcal{R}}$ each of which is the interpretation of $f_j \in \mathcal{F}$.[34]

For example, the structures associated to the language of groups are just the groups themselves. In each group, the identity element symbol is interpreted as the identity element and the group operation symbol is interpreted as the group operation.

In every structure, the language of that structure induces a collection of subsets on that structure which we call the *definable sets*. Intuitively, these are the sets that we can talk about in our language. A more precise definition is given here:

Definition 3.5.3. Let $\mathcal{R} = (R, ...)$ be an $\mathcal{L}$-structure. We say that $X \subseteq R^n$ is **definable** if there is an $\mathcal{L}$-formula $\varphi(x_1, ..., x_n, y_1, ..., y_m)$, and parameters $p_1, ..., p_m$ such that $X = \{(a_1, ..., a_n) \in R^n : \mathcal{R} \models \varphi(a_1, ..., a_n, p_1, ..., p_n)\}$. We say that φ **defines** X.[35]

For example, in the group of integers, the set of even integers is definable by the formula $\varphi(x) := \exists y(y + y = x)$. But many subsets of structures cannot be defined in the language of the structure. For example, the set of real numbers cannot be defined in the field of complex numbers. It is usually not obvious which subsets of a given structure are definable.

The structures that we call *o-minimal* are defined by the intuitive simplicity of their definable sets. In all o-minimal structures, the definable subsets are *finite* unions of isolated points and intervals. More precisely:

Definition 3.5.4. An $\mathcal{L}$-structure $\mathcal{R} = (R, <^{\mathcal{R}}, ...)$, where $<^{\mathcal{R}}$ is a dense linear order without endpoints, is called **o-minimal** if and only if every set $S \subseteq R^1$ that is definable in $\mathcal{R}$ with parameters is a union of finitely many intervals and points.[36]

The two simplest kinds of o-minimal structures are ordered vector spaces of the form, $(R, <^{\mathcal{R}}, +^{\mathcal{R}}, (\lambda \cdot^{\mathcal{R}})_{\lambda \in F}, 0^{\mathcal{R}})$, and real closed fields of the form, $(R, <^{\mathcal{R}}, +^{\mathcal{R}}, \cdot^{\mathcal{R}}, 0^{\mathcal{R}}, 1^{\mathcal{R}})$. The

[34] Dries (1998), p 21.

[35] Marker (2002b), p 19. We also say that $\mathcal{L}$ *generates* the sequence of definable sets $\delta = (\{X \subseteq R^i : X \text{ is definable}\})_{i \in \mathbb{N}}$. (Dries 1998, p 21)

[36] Dries (1998), p 23.

smallest o-minimal ordered vector space is $(\mathbb{Q}, <^{\mathbb{Q}}, +^{\mathbb{Q}}, (\lambda \cdot^{\mathbb{Q}})_{\lambda \in \mathbb{Z}}, 0)$, and the smallest real closed field is $(\mathbb{A}, <^{\mathbb{A}}, +^{\mathbb{A}}, \cdot^{\mathbb{A}}, 0, 1)$.[37]

But the above definition of o-minimal structures also *rules out* many structures. For example, if we add the sine function symbol to our language, the resulting structures (over dense linear orders) will not be o-minimal. This is because we can use the sine function symbol to define the set of integers. (Simply let $\varphi(x)$ be the formula, $\sin(x) = 0$.) And since the integers cannot be written as a finite union of points and intervals, this structure is not o-minimal.

O-minimal structures have a number of nice logical and geometrical properties. In particular, we will see that the *o-minimal continuous functions* within every o-minimal structure are 'visualizable' in precisely the ways that the ϵ-δ functions were not. Intuitively, an o-minimal continuous function is any topologically continuous function that the language has resources to talk about. More precisely,

Definition 3.5.5. A function $f : R \to R$ in an o-minimal structure is **o-minimal continuous** if and only if the inverse image of every definable open set of R is definable and open.[38]

For example, in the field of real numbers, even though $\sin(x)$ is a continuous function, it is not *o-minimal* continuous because the language of fields does not have the resources to talk about this function.

In every language, we can collect together sentences of the language to form a *theory*. If all of the sentences of a theory T are true on a structure $\mathcal{M}$, we say that the structure *models* that theory. We express this symbolically as: $\mathcal{M} \models T$. If all the models of a given theory are o-minimal, we say that the theory itself is o-minimal. That is:

[37] Here, '$\mathbb{A}$' refers to the field of real algebraic numbers, whose underlying set is the set of real roots of the non-zero polynomials in one variable with rational coefficients.

[38] My term "o-minimal continuous" is synonymous with the Dries' (1998) term "definable continuous." Also, the topology of an o-minimal structure is always the order topology.

Definition 3.5.6. An $\mathcal{L}$-theory T is **o-minimal** iff for all models $\mathcal{M} = (M, ...)$, such that $\mathcal{M} \models T$, the definable sets of M^1 are exactly the finite unions of points and intervals.

For example, the theory of real closed fields is an o-minimal theory. And so is the theory of dense linear orders. In each case, the axioms of the theory prohibit all their models from having definable subsets that contain infinitely many isolated points. That is, all their models are o-minimal.

Let us close this overview of o-minimality with the definition of an *o-minimal expansion.* Intuitively, this is an o-minimal structure that we attain by enriching the language of a different o-minimal structure. More precisely:

Definition 3.5.7. Let $\mathcal{L}^+ \supset \mathcal{L}$. If $\mathcal{M}^+$ is an $\mathcal{L}^+$-structure, then by ignoring the interpretations of the symbols in $\mathcal{L}^+ \setminus \mathcal{L}$ we get an $\mathcal{L}$-structure $\mathcal{M}$. We call $\mathcal{M}$ a **reduct** of $\mathcal{M}^+$ and $\mathcal{M}^+$ an **expansion** of $\mathcal{M}$. An **o-minimal expansion** of an o-minimal structure, is one that preserves o-minimality (i.e. one in which the new function and relation symbols cannot be used to define new subsets of R^1).[39]

The notion of an o-minimal expansion allows us to state one of the first important results about the class of o-minimal structures. Peterzil and Starchenko (1998) have proven that all non-trivial o-minimal structures are expansions of these two basic kinds of structures.[40] That is, they are all either (1) order vector spaces of the form $(R, <^{\mathcal{R}}, +^{\mathcal{R}}, (\lambda \cdot^{\mathcal{R}})_{\lambda \in F}, 0^{\mathcal{R}}, \mathcal{F})$, where $\mathcal{F}$ stands for the new functions of the expansion, or (2) real closed fields of the form $(R, <^{\mathcal{R}}, +^{\mathcal{R}}, \cdot^{\mathcal{R}}, 0^{\mathcal{R}}, 1^{\mathcal{R}}, \mathcal{F})$. The most well-known expansion of a real closed field is Wilkie's real exponential field (i.e. $R = \mathbb{R}$ and $\mathcal{F} = \{e^x\}$).[41]

The class of o-minimal structures has many other notable properties. From the perspective of pure model theory, the two most important are the following. First, all o-minimal

[39] Marker (2002b), p 31.

[40] "Trivial" is a technical term for Peterzil and Starchenko. Roughly, it means that the number of definable subsets of the structure is 'small'.

[41] Wilkie (1996). This result also seems to have led to a significant increase in the number of mathematicians working on o-minimality.

theories have the Non Independence Property.[42] Second, that o-minimality is preserved under elementary equivalence of structures.[43] Results like these have suggested that the class o-minimal structures is unified enough to be worthy of sustained mathematical inquiry.

We are now in a position to state the o-minimal IVT and turn back to the question of whether it has a valid connection with visual imagination. Here is the o-minimal formulation of the Intermediate Value Theorem, and its well-known proof:

> (OI) Let $\mathcal{R}$ be an o-minimal structure with a constant symbol '0'.[44] If the function $f : [a, b] \subseteq R \to R$ is o-minimal continuous and $f(a) < 0 < f(b)$, then there is a c such that $f(c) = 0$.[45]

> **Proof:** Suppose not. Then there is no x such that $f(x) = 0$. In other words, $[a, b] = f^{-1}((-\infty, 0)) \cup f^{-1}((0, \infty))$.[46]
>
> Since $\mathcal{R}$ is o-minimal, the definable subsets of R^1 are exactly the finite unions of points and intervals. Hence, the subsets $(-\infty, 0)$ and $(0, \infty)$ are definable. Furthermore, since R is o-minimal, it has the order topology, so $(-\infty, 0)$ and $(0, \infty)$ are open, too. Now, since f is definable continuous, the inverse images of definable open sets are definable and open. Hence, $f^{-1}((-\infty, 0))$ and $f^{-1}((0, \infty))$ are definable and open. Because they are both definable, they can each be written as a finite disjoint union of open intervals in the subspace $[a, b]$ of R. That is, we can write:
>
> $f^{-1}((-\infty, 0)) = [a, c_1) \cup (c_2, c_3) \cup ... \cup (c_{m-1}, c_m)$, where $a < c_i < c_{i+1} < b$ for all $i < m$.
>
> $f^{-1}((0, \infty)) = (c'_1, c'_2) \cup (c'_3, c'_4) \cup ... \cup (c'_n, b]$, where $a < c'_j < c'_{j+1} < b$ for all $j < n$.
>
> For contradiction, consider the endpoint, c_1, of the first interval of $f^{-1}((-\infty, 0))$. Since $[a, b]$ is the disjoint union of the above two sets, we know that c_1 is in

[42] This notion was developed by Shelah in the 1960s and 1970s. The distinction between theories that have the Independence Property and those that have NIP is widely considered to be "one of the most important dividing lines between theories" (Peterzil 2007, p 13). For more on Stability Theory and its basic notions, see Buechler (2017).

[43] Peterzil (2007), §3. This result is important because the related property of being 'minimal' (i.e. all definable subsets are either finite or co-finite) is *not* preserved under elementary equivalence.

[44] The restriction to languages with a zero symbol is not strictly necessary.

[45] Dries (1998), p 19. For an additional statement of the proof, see Peterzil (2007).

[46] $f^{-1}((-\infty, 0)) := \{x \mid f(a) \leq f(x) < 0\}$, and $f^{-1}((0, \infty)) := \{x \mid 0 < f(x) \leq f(b)\}$.

exactly one of these sets. By stipulation, it cannot be in $f^{-1}((-\infty, 0))$ because its intervals are disjoint and open. So $c_1 \in f^{-1}((0, \infty))$. This means that there is some j such that $c_1 \in (c'_j, c'_{j+1})$. But, since (c'_j, c'_{j+1}) is open, there is a point $q < c_1$ such that $q \in (c'_j, c'_{j+1})$. And since $a < c'_j$ by definition, it must be the case that $q \in [a, c_1)$ as well. From this, it follows that q is a member of both $f^{-1}((-\infty, 0))$ and $f^{-1}((0, \infty))$. Hence, $f(q) < 0$ and $f(q) > 0$. $\rightarrow\leftarrow$ $\square$

The above proof provides us with a certain kind of mathematical knowledge of the o-minimal IVT, a form of knowledge based on *justification*. This is not the kind of knowledge that I am primarily concerned with in this paper. The kind of knowledge I am interested in is *imagery-based mathematical entitlement*, which is not acquired through explicit proof. In order to show that the o-minimal IVT *also* admits of imagery-based entitlement, we will take note of three relevant theorems from o-minimality, namely the Monotonicity Theorem, the Piecewise-Differentiability Theorem, and Dimension Invariance Theorem. These are stated as follows:

Monotonicity Let $\mathcal{R} = (R, <, \ldots)$ be an o-minimal structure, and let $f :]a, b[] \to R$ be a definable function on some open interval $]a, b[\subseteq R$. Then there are $a_0 = a < a_1 < \ldots < a_n = b$ such that on each $]a_i, a_{i+1}[$ the function f is either constant or strictly monotone and continuous.[47]

Piecewise-Differentiability Let $\mathcal{R} = (R, <, \ldots)$ be an o-minimal structure, and let $f :]a, b[\to R$ be a definable function on some open interval $]a, b[\subseteq R$. Then for every $n \in \mathbb{N}$ there exist $a = a_0 < a_1 < \ldots < a_r = b$ such that f is n-differentiable on each $]a_i, a_{i+1}[$.[48]

Dimension-Invariance If $X \subseteq R^m$ and $Y \subseteq R^n$ are definable subsets in an o-minimal structure and there is a definable bijection between X and Y, then $\dim X = \dim Y$.[49]

Together these three theorems tell us that o-minimal continuous functions lack the sort of pathological behavior exhibited by ϵ-δ continuous functions. The topologist's sine curve is

[47] Peterzil (2007), p 4.
[48] Peterzil (2007), p 19.
[49] Dries (1998), p 64. Proposition 1.3.

not an o-minimal continuous function because it violates Monotonicity. Similarly, the Weierstrass function violates Piecewise-Differentiability and the Hilbert curve violates Dimension-Invariance. These three theorems do more than rule out a few bad apples. They establish that every o-minimal continuous function, when plotted in a Cartesian reference frame, conforms to the category specification of the concept *planar curve*: (i) bounded contour with no visible gaps, (ii) piecewise monotonic, (iii) piecewise smooth, and (iv) passes the vertical line test. I would like to suggest that the concept *o-minimal continuous function* has, if not the the *same* category specification as the concept *planar curve*, something very similar to it. Of course, these two concepts themselves remain very much distinct. Possessing the concept *o-minimal continuous function* requires that one accepts a number of claims and inferences about model-theoretic entities. The same cannot be said for the concept *planar curve*. Nevertheless, every object falling under these two concepts can be represented using the limited representational resources available in visual imagination. This was not true for the concept *ϵ-δ continuous function*. Whatever these limited resources turn out to be, I think the claim they can represent all planar curves but not all ϵ-δ continuous functions is fairly secure. I believe this claim is all we need to establish that the concept *o-minimal continuous function* is capable of having a category specification.[50]

Supposing that the concept *o-minimal continuous function* has the above category specification, is there reason to think that the o-minimal IVT has a valid connection to visual imagination? Yes, because we can write this valid connection in argument form just as we did for (SI):

<u>Valid Connection for (OI)</u>

P1. All o-minimal continuous functions $f(x)$ are piecewise smooth and monotonic. (Specification for *o-minimal continuous function*)

[50] In Part 6, I will argue for the stronger claim that the developers of o-minimality *intended* for these o-minimal concepts to have category specifications.

P2. All piecewise smooth and monotonic contours from below the x-axis to above have a segment that is smooth, monotonic and goes from below the x-axis to above. (Structure of visual imagination)

P3. All smooth, monotonic curve segments that go from below the x-axis to above it will *intersect* the x-axis at a point. (Structure of visual imagination)

P4. For all curves that intersect the x-axis, there is a $c \in R$ such that $f(c) = 0$. (Concept of *o-minimal function*[51])

C. Therefore, for any o-minimal continuous function $f(x)$ that goes from below the x-axis to above it, there is a $c \in R$ such that $f(c) = 0$.

This argument establishes the existence of a valid connection between (OI) and visual imagination. Recall that no such connection could be established for the case of (AI) because the concept ϵ-δ *continuous function* did not have a category specification. In the case of (OI), by contrast, the premise P1 can be derived from the category specification of the concept *o-minimal continuous function*.

In this section, I have shown that the concept *o-minimal continuous function* is capable of having a category specification and that there is a valid connection between the o-minimal IVT and visual imagination. In Part 6, I will establish the stronger claim that the o-minimal IVT can be known through imagery-based entitlement. In order to establish the stronger claim I will show that o-minimal concepts, such as *o-minimal continuous function*, do indeed have category specifications.

[51] This claim should be primitively compelling to everyone who possesses the concept *o-minimal function*.

3.6 O-minimality as a Visual-Epistemological Project

3.6.1 Can o-minimal objects be visualized?

There is good reason to think that the concept *o-minimal continuous function* has a category specification. This should not be obvious. It could be that o-minimality, like abstract algebra, is not really concerned about whether its concepts have category specifications. (Consider, for example, the abstract-algebraic concept *algebraically independent set.*) I will argue for this claim on the grounds that the developers of o-minimality *intended* its subject matter to be inherently visualizable. Along the way I will suggest that these mathematicians are engaged in a mathematical project that is grounded in deep epistemological commitments about the foundations of mathematics. Unfortunately, I will not be able to evaluate the validity of these commitments here. Instead, I will simply make note of these commitments when they bear on the imagistic nature of o-minimal objects.

The first textbook on o-minimality, *Tame Topology and O-minimal Structures*, was published in 1998 by Lou van den Dries. He glosses the subject matter of o-minimality as "the realm of geometry and topology envisaged by Poincaré" (Dries 1998, p 1). The main work in which Henri Poincaré put forward his views on topology and geometry was his "Analysis Situs" papers.[52] In the introduction to this work, Poincaré tells us what he takes to be the subject matter of geometry:

[52]This paper is widely considered to be a founding document of algebraic topology: "The systematic study of algebraic topology was initiated by the French mathematician Henri Poincaré (1854-1912) in a series of papers during the years 1895-1901. Algebraic topology, or analysis situs, did not develop as a branch of point-set topology. Poincare's original paper predated Frechet's introduction of general metric spaces by eleven years and Hausdorff's classic treatise on point-set topology, *Grundzüge der Mengenlehre*, by seventeen years. Moreover, the motivations behind the two subjects were different. Point-set topology developed as a general, abstract theory to deal with continuous functions in a wide variety of settings. Algebraic topology was motivated by specific geometric problems involving paths, surfaces, and geometry in Euclidean spaces. Unlike point-set topology, algebraic topology was not an outgrowth of Cantor's general theory of sets. Indeed, in an address to the International Mathematical Congress of 1908, Poincaré referred to point-set theory as a "disease" from which future generations would recover." (Croom 1978, p 2).

> Geometry, in fact, has a unique *raison d'être* as the immediate description of the structures which underlie our senses. (Poincaré 1895, p 5)

From this characterization, it is clear that o-minimality, insofar as it shares its subject matter with the geometry of Poincaré, is a field of study explicitly tailored to visual perception and visual imagination.

Lou van den Dries also claims that o-minimality is "an excellent framework for developing tame topology, or *topologie modérée*, as outlined in Grothendieck's prophetic 'Equisse d'un Programme' of 1984" (Dries 1998, vii). A few other contemporary model theorists have expressed the same view.[53] Grothendieck characterizes his theory of 'tame topology' as an alternative to general topology (i.e. the axioms of a topological space) as a foundation for the mathematical field of geometric topology.[54] His "heuristic reflections" on the foundations of topology reveal that he thinks the originators of the general-topological treatment of geometric topology started on the wrong foot:

> 'General topology' was developed (during the thirties and forties) by analysts and in order to meet the needs of analysis, not for topology per se, i.e. the study of the topological properties of the various geometrical shapes. That the foundations of topology are inadequate is manifest from the very beginning, in the form of "false problems" (at least from the point of view of the topological intuition of shapes) such as the "invariance of domains" ... (Grothendieck 1984, pp 28-29)

What is remarkable about this document is that Grothendieck is calling for a deep revision in the foundations of general topology, not (primarily) because his proposed alter-

[53] For instance, see Marker (2002a), Wilkie (2007), and Yomdin and Comte (2004).

[54] Grothendieck (1984, note 6) suggests that he wants tame topology to provide a restatement and proof of the *Hauptvermutung* (i.e. any two triangulations of a triangulable space have a common refinement, a single triangulation that is a subdivision of both of them), which is false in the received, general-topological foundation of geometric topology.

native has greater problem-solving capabilities (which I take to be, historically, the rule in shifts of mathematical norms), but because general topology is 'unnatural' given its task: to provide a framework for the "topological [and] geometrical intuition of shapes".[55] Indeed, he even thinks that his revisions of general topology go hand-in-hand with the need to "rewrite a new version, in modern style, of Klein's classic book on the icosahedron and the other Pythagorean polyhedra".[56] The objects that would be studied in such a book, although written in the notations of algebraic geometry, are presented to us visually:

> I do not believe that a mathematical fact has ever struck me quite so strongly as [a certain theorem of algebraic geometry about complex algebraic curves], nor had a comparable psychological impact. This is surely because of the very familiar, non-technical nature of the objects considered, of which any child's drawing scrawled on a bit of paper (at least if the drawing is made without lifting the pencil) gives a perfectly explicit example. (Grothendieck 1984, p 247)

These facts suggest that o-minimality, insofar as it shares its subject matter with Grothendieck's tame topology and Poincaré's geometry, is explicitly concerned with the visualizable realm.

A final reason to think that o-minimality is inherently visualizable arises when we consider the origins of o-minimality. It is well known that o-minimality grew out of work by Alfred Tarski on the method of quantifier elimination in geometry (Tarski 1948). Tarski applied this method to his axiomatization of elementary geometry. His axiom system for geometry attempted to capture as many of the results of Euclid's *Elements* as he thought feasible (Tarski 1959). And since the *Elements* is an essentially diagrammatic treatise, it would seem that Tarski's subject matter, at least in the 1959 paper, is inherently visualizable.

[55] Grothendieck (1984), pp 240, 258, 259.

[56] Grothendieck (1984), p 255. The book that Grothendieck mentions is Felix Klein's *Lectures on the Icosahedron* (1884).

This hypothesis is confirmed when Tarski says that in the formal language of elementary geometry,

> we are able to ... refer ... to various special classes of geometrical figures, such as the straight lines, the circles, the segments, the triangles, the quadrangles, and, more generally, the polygons with a fixed number of vertices... This is primarily a consequence of the fact that, in each of the classes just mentioned, every geometrical figure is determined by a fixed finite number of points. (Tarski 1959, pp 16-17)

By Tarski's own lights, his axiomatization of geometry has circles, polygons, lines, etc., as its subject matter. These are paradigm cases of mathematical objects that can also form the contents of visual imagination. We cannot overlook the fact that van den Dries (1984, p 99) motivates the definition of an o-minimal structure by appealing to the "geometric idea" in Tarski's method of quantifier elimination.[57] Finally, the fact that Tarski's geometric structures are all real closed fields further confirms the idea that o-minimality and Tarski's geometry share a substantial part of their subject matter. These considerations about Tarski, support my claim that the primary objects of o-minimality can be represented in visual imagination.

From the above discussion of Poincaré, Grothendieck, and Tarski, it is clear that the developers of o-minimality see its subject matter as inherently visual. Therefore, it is reasonable to think that the concept *o-minimal continuous function* has a category specification. Furthermore, the category specification of *o-minimal continuous function* given by the description set in §5.1 presents itself as an obvious candidate. It is an obvious candidate

[57] The "geometric idea" of Tarski's method of quantifier elimination contrasts with the "decidability idea" of Tarski, which is a purely computational idea. In fact, van den Dries (1984, p 98) says that the search for a decision method is a "waste of time" in the case of the elementary of theory of the real exponential field. The main results of van den Dries' 1984 paper are a collection of "finiteness theorems" that one might take to be essential to our pre-theoretic notion of geometricality (i.e. what Giaquinto calls our "visual expectations in mathematics" (Giaquinto 2007, p 3). These results will be discussed below.

because o-minimal continuous functions demonstrably satisfy each of the descriptions in §5.1: no gaps, piecewise monotonic, piecewise smooth, and satisfies the vertical line test.[58]

I take the above argument to support my claim that the o-minimal IVT admits of imagery-based mathematical entitlement.[59] In the final section, I will discuss the extent to which o-minimality has succeeded in developing a general theory for which our visual faculties are reliable guides to the behavior of its primary objects. I will focus on how the visualizable objects of o-minimality behave differently from the real-analytic objects that led to the so-called 'flight from intuition' in the nineteenth century. I will argue that o-minimality has succeed and the behavior of o-minimal visualizable objects, unlike the behavior of visualizable objects in real analysis, "accords with our visual comprehension of space and spatial objects" (Giaquinto 2007, p 5).

3.6.2 Visual Virtues of O-minimality

Some of the most important results in o-minimality are what one might call 'tameness results' (Wilkie 2007, Marker 2002a). These tameness results suggest that o-minimality has succeed in isolating a subject matter that can be reliably investigated with one's visual faculties. The tameness results of o-minimality can, very roughly, be split into two kinds: invariance results and finiteness results.[60] In the former category belong the propositions that Dimension and Euler Characteristic are invariant across definable bijections. In the latter category, there are a number of 'finiteness' theorems. We have already seen two

[58]Moreover, there is a "general procedure" for imagining the o-minimal continuous functions, since each has a finite number of smooth pieces. According to Giaquinto's (2011, pp 299-300) psychological hypothesis, this disposes us to generalize over all o-minimal continuous functions.

[59]Although I cannot defend it here, I think parallel arguments can be made for the reality of imagery-based entitlement in the case of other o-minimal theorems, such as Rolle's Theorem and the Mean Value Theorem.

[60]There are a few theorems that do not fit into either of these categories: the Curve Selection Theorem, the Definable Choice Theorem, the Trivialization Theorem, and the theorem that states that all definable groups are definably isomorphic to a Lie group. See Peterzil (2007) and Marker (2002a). I leave to future investigation the question of to what extent these additional theorems suggest that o-minimal structures are 'tame'.

of these theorems, the Monotonicity Theorem and the Piecewise-Differentiability Theorem, which together state that all definable continuous functions can be decomposed into a *finite* number of pieces, each of which is either monotonic and n-differentiable. But there are many other finiteness theorems. These finiteness theorems show that the developers of o-minimality succeeded in their attempt to delineate a whole collection of concepts with category specifications. In other words, our visual faculties are reliable guides to the behavior of many o-minimal objects, not merely o-minimal continuous functions.

I believe that these finiteness theorems can be used to provide a clear distinction between o-minimality and real analysis. According to Giaquinto, the main reason why real analysis eludes visual imagination is the non-finite nature of its objects:

> The reliability of visual thinking in mathematics, especially in analysis, came under heavy suspicion in the nineteenth century. The main reason was that our visual expectations in mathematics, known collectively as geometrical or spatial intuition, quite often turned out to be utterly misleading, particularly about what happens "at the limit" of an infinite process. A prominent case is the existence of ...continuous but nowhere-differentiable function[s] ... A related example reinforced suspicion. In 1890 Peano showed that it is possible to define a curve that completely fills a two-dimensional region ... (Giaquinto 2007, pp 3-4).

There is no question that our intuitions about the behavior of infinite totalities and the results of infinite processes can be very misleading. But o-minimality, I claim, is specifically suited to avoid these kinds of objects and processes. There are three reasons o-minimality can avoid these misleading infinite totalities and processes: the definition of an o-minimal structure, the finiteness theorems hinted at above, and the fact that dimension is a definable invariant. I will discuss each of these below.

The *definition* of an o-minimal structure guarantees that we will never find ourselves talking about one-dimensional objects that do not have finitely many connected components. In other words, we can only define subsets of R^1 that are finite collections of points, line segments, and lines. We can express the content of this claim in terms of the limiting behavior of these definable sets. That is, the definition of an o-minimal structure guarantees that, as we approach a point in R^1, we will not find any definable set D that infinitely oscillates between $p \in D$ and $p \notin D$. In fact, expression of infinite oscillation are familiar from tense logic and relation modal logic.[61] It is therefore natural to ask, given the order topology on all o-minimal structures, whether we can formalize the above intuitive condition on infinite oscillations in *topological modal semantics*.[62]

In the topological modal semantics, '$\Box p$' says that there is an open neighborhood around every point such that p is true on that neighborhood. Within this formal framework we can express the claim that there are infinite oscillations between p and $\neg p$ around a single point x by the following holding at x: $\neg\Box(p \to \Box p) \wedge \neg\Box(\neg p \to \Box\neg p)$. But I have shown that this can never happen on o-minimal expansions of the real numbers. Indeed the negation of the above sentence is characterist of all such expansion. See Appendix for details. This theorem provides us with a formal statement of the intuitive fact that the one-dimensional definable subsets are finite unions of familiar, visualizable objects like points and line segments.

Let us move on to our discussion of the finiteness theorems. These theorems give us yet another reason to think o-minimality can avoid the misleading infinite totalities and processes that afflicted 19th-century analysis. Above, we said that the Piecewise-Differentiability Theorem guarantees that all definable continuous functions will be n-differentiable at all but finitely many points. Hence, we avoid continuous nowhere-differentiable functions. We

[61] For example, in tense logic, the formula '$\mathrm{G}(\mathrm{F}p \wedge \mathrm{F}\neg p)$' expresses the fact that at any future time there will be a time in the future when p is true and a time in the future when p is false.

[62] Topological modal semantics originated with Tarski's 1938 paper, "Sentential Calculus and Topology". More recently, the study of Spatial Logics has been focused on the topological modal semantics.

also said that the Monotonicity Theorem guarantees that all definable continuous functions will only have finite oscillations. Hence, we avoid paradoxes of infinitely oscillating curves like the topologist's sine curve. Other pathological and 'untame' constructions are ruled out by additional finiteness theorems like the Finite Homeomorphism Type Theorem, the Cell Decomposition Theorem, the Stratification Proposition, and the Triangulation Theorem. These finiteness theorems provide the basis for additional suggestive results and conjectures such as the following:

- The Borel subsets of $\mathbb{R}^n$ are not definable. (Marker 2002a, p 353)
- The solutions to definable differential equations have a finite number limit cycles. (Dries and Speissegger 2003, p 2)
- O-minimal theories do not exhibit the "Gödel Incompleteness Phenomena". (Steinhorn 2003, p 21; Haskell, *et al.* 2000, p 7)

It is difficult to articulate what, if any, visual consequences these statements have. Nevertheless, they are concrete cases in which the objects of o-minimality exemplify a finite or tame behavior that simply does not hold for objects in real analysis.

Finally, the invariance theorems of o-minimaility guarantee that visually detectable properties like *dimension* and *euler characteristic* will be preserved under definable homeomorphism. In contrast, the space-filling curves and other non-visual equivalences of real analysis reveal that these objects do not align with our spatial expectations of the behavior of a geometric figure when it is deformed. After all, it would seem quite puzzling if a line segment could fill a square simply by deformation. In fact, Wilkie describes dimension invariance as a *desideratum* of o-minimality:

> [Any framework for tame topology] should have built-in restrictions so that we are *a priori* guaranteed that pathological phenomena can never arise. In particular,

> there should be a meaningful notion of dimension for all sets under consideration and any that can be constructed from these ... (Wilkie 2007, p 1)

Given the above discussion of the definition, finiteness theorems, and invariants of o-minimality, I believe there is strong support for the claim o-minimality has succeeded in developing a theory for which our visual faculties are reliable guides to the behavior of its primary objects.

Before leaving this topic of the visual virtues of o-minimality, I would like to mention that o-minimality also allows us to reject a dichotomy that Giaquinto puts to us when he says that we have to choose between the IVT on the one hand, and the PDT and Dimension-invariance on the other:

> [The] maintain[enance of] our cognitive predispositions about space motivates the intermediate-value condition on continuous functions, as one of our strongest cognitive predispositions is that space has no gaps, not even invisibly small gaps.
>
> Of course some things had to give: our cognitive dispositions are violated in classical analysis by the existence of continuous nowhere-differentiable functions and space-filling curves. We can avoid these surprises by replacing classical analysis with smooth infinitesimal analysis. But the IVT does not hold in smooth infinitesimal analysis. (Giaquinto 2011, p 303.)

We do not have to choose between the IVT on the one hand, and the PDT and dimension-invariance on the other hand. In o-minimality, we can maintain all three of these expectations (along with many others). Moreover, the o-minimal IVT does not require extending the rationals to reals. We can get away with incomplete orders because the proof of the o-minimal IVT never appeals to the existence of minima or suprema, which stands as a counterexample to what Giaquinto says about the IVT here:

> [I]f the IVT is a *desideratum* for an account of the real numbers, it has to be proved for the given account, and that cannot be done without appeal to Dedekind Completeness or something which implies it. (2011, p 303) [63]

In sum, o-minimality seems tailored to our cognitive expectations about the spatial contents of visual imagination. The definitions, finiteness theorems, and invariance theorems suggest that o-minimality is able to sidestep much of the misleading 'limiting' behavior that plague the basic objects of real analysis.

In this essay, I have articulated a general account of imagery-based entitlement and applied it to the the case of the o-minimal Intermediate Value Theorem. More generally, I have argued that we have good reason to think the subject matter of o-minimality is inherently visual. Finally, I proposed that the differential epistemic status of theorems in real analysis and and theorems in o-minimality can be explained by considering the definitions, finiteness theorems, and invariance results that hold in o-minimality but not in real analysis. I believe my arguments in this essay undermine any attempt to restrict imagery-based mathematical entitlement to elementary geometry and arithmetic. If this is correct, a comprehensive understanding of mathematical knowledge will have to go beyond the study of proof-based justification. I would like to think the most promising way to study non-inferential mathematical knowledge, such as imagery-based entitlement, is by synthesizing results in cognitive psychology with a detailed study of how mathematicians understand their subject matter. Throughout this essay, I hope to have exemplified this methodological commitment.

[63] For further discussion of the inessential nature of Dedekind Completeness for the IVT in real closed fields, see Sinaceur (1994).

3.7 Appendix: Modal Logic of O-minimal Structures

I will show here that the class of o-minimal structures can be characterized in propositional modal logic with the definable topological semantics.[64] First, I say what the definable topological semantics is, and then I prove the theorem stated in the main text.

Definition 3.7.1. Let $\mathcal{R} = (\mathbb{R}, <, +, \times, \ldots)$ be an o-minimal expansion of the real field. In the **definable topological semantics**, we let a definable valuation $V : SL \to \mathcal{D}$, be a map from atomic letters to the collection of definable-with-parameters subsets of $\mathbb{R}$, such that for all p and x[65]:

$$V, \mathcal{R}, x \models p \iff x \in V(p) \tag{3.1}$$
$$V, \mathcal{R}, x \models \Box\varphi \iff \text{There is open } U \ni x \text{ such that } U \subseteq \{y \in \mathbb{R}^k : V, \mathcal{R}, y \models \varphi\} \tag{3.2}$$

The clauses for the propositional connectives are the usual ones, and here the topology is the usual Euclidean topology on $\mathbb{R}^k$. Further, as usual, one says that $\mathcal{R}, x \models p$ if and only if for all definable valuations V, we have $V, \mathcal{R}, x \models p$.

—

Theorem 1. For any expansion of the real field $\mathcal{R}$, $\mathcal{R}$ is o-minimal iff for all $x \in \mathbb{R}$, and p,

$$\mathcal{R}, x \models \Box(p \to \Box p) \vee \Box(\neg p \to \Box\neg p) \tag{3.3}$$

Proof. Note that $\Box(p \to \Box p) \vee \Box(\neg p \to \Box\neg p)$ is logically equivalent to $S := \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$. In this proof, we will be using the second formula, S.

'$\Rightarrow$': Let $\mathcal{R}$ be o-minimal. Let D_i be an arbitrary definable subset of $\mathbb{R}$. Then D_i is a finite union of isolated points and intervals. Let V_i be the valuation such that $V_i(p) = D_i$. We must show for an arbitrary $x \in \mathbb{R}$, that $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$. There are two cases. Case (I): $V_i, \mathcal{R}, x \models p$. Case (II): $V_i, \mathcal{R}, x \models \neg p$.

[64] Of course there are many other kinds of "topological semantics" for propositional modal logic. Some alternative topological interpretations of the Box operator that are worthy of consideration include: the derivate operator, the definable closure operator, and the convex closure operator. All of these have been studied in some capacity before, but (as far as I know) no one has studied them in the o-minimal setting.

[65] The following clauses are adapted from van Benthem and Bezhanishvili (2007), p 218.

In case (I), I claim that $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$. That is, there is an open interval $O_x \ni x$ such that for all $y \in O_x$, we have $V_i, \mathcal{R}, y \models p \vee \Box\neg p$. Since D_i is a finite union of isolated points and (open or closed) intervals, either (a) x is an isolated point in D_i, or (b) x is in an (open or closed) interval $U_x \subseteq D_i$. Consider (a). If x is an isolated point in D_i, then there is an open interval J_x around x such that $J_x \cap D_i = \{x\}$. Let $I := J_x$. Then for all $y \in I$, if $x = y$, then $V_i, \mathcal{R}, y \models p$, and if $x \neq y$, then $V_i, \mathcal{R}, y \models \Box\neg p$. Hence, $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$. Consider (b). Since there is an interval U_x containing x, either: (b1) $x \in Int(U_x)$, (b2) x is an endpoint of U_x. Consider (b1) and let $I := Int(U_x)$. Then for all $y \in I$, we have $V_i, \mathcal{R}, y \models p$. Hence, $V_i, \mathcal{R}, x \models \Box p$, entailing that $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$. Consider (b2). Since x is an endpoint of U_x, there is a $w \notin D_i$ such that, for all points z between x and w (without loss of generality we can assume $x < w$), $V_i, \mathcal{R}, z \models \neg p$ (and, thus also: for all z, we have $V_i, \mathcal{R}, z \models \Box\neg p$). Let $I := Int(U_x) \cup [x, w[$. Then for all $y \in I$, either $y \in U_x$ (and so $V_i, \mathcal{R}, y \models p$), or $y \in]x, w[$ (and so $V_i, \mathcal{R}, y \models \Box\neg p$). Hence, $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$.

In case (II), I claim that $V_i, \mathcal{R}, x \models (\neg p \vee \Box p)$. Here we follow the same argument as in case (I), except we consider the complement of D_i, $\mathbb{R} \setminus D_i$.

Since we began with an arbitrary D_i and an arbitrary x, we can generalize over all worlds x and definable valuations V_i. We thus conclude, for any $x \in \mathbb{R}$, that $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$.

'$\Leftarrow$': Let $\mathcal{R}$ be any expansion of the real field in which, for all $x \in \mathbb{R}$ and all $D_i = V_i(p)$, $V_i, \mathcal{R}, x \models \Box(\neg p \vee \Box p) \vee \Box(p \vee \Box\neg p)$. We need to show that $\mathcal{R}$ is o-minimal. Appealing to Schoutens (2014), Corollary 2.4, we have that if $\mathcal{R}$ is locally o-minimal, then $\mathcal{R}$ is o-minimal.[66] Hence, all we need to show is that $\mathcal{R}$ is locally o-minimal. By Toffalori and Vozoris (2009), Proposition 2.4:

> $\mathcal{R}$ is locally o-minimal iff for every $x \in \mathbb{R}$ and every $D_i \subseteq \mathbb{R}$, there are $c, d \in \mathbb{R}$ such that $c < x < d$ and either $D_i \cap]c, d[$ or $]c, d[\setminus D_i$ is equal to one of the following: (i) $\{x\}$, (ii) $]c, x]$, (iii) $[x, d[$, or (iv) the whole interval $]c, d[$.

We will show that if every x and V_i is such that $V_i, \mathcal{M}, x \models \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$, then for each x, there is an open interval I around x such that for all D_i, either $D_i \cap I$ or $I \setminus D_i$ is equal to any of (i)-(iv) above. We can then conclude that $\mathcal{R}$ is o-minimal. Fix x and D_i. In order to show that either $I \cap D_i$ or $I \setminus D_i$ is equal to (i)-(iv), we will start by considering the consequences of our supposition under the valuation V_i. That is, we know from our supposition that $V_i, \mathcal{R}, x \models S$. We thus know that one of two cases holds. Case (I), $V_i, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$. Case (II), $V_i, \mathcal{R}, x \models \Box(\neg p \vee \Box p)$.

In case (I), there is an open interval O_x around x such that for all $y \in O_x$, $V_i, \mathcal{R}, y \models p \vee \Box\neg p$. The second disjunct tells us that $O_x \setminus D_i$ is open. By Folland (1999), Proposition

[66]Schoutens actually proves that if $\mathcal{R}$ is type-complete, then it is o-minimal. But, as Rennet (2014), p 55, fn 1, points out this type-completeness is equivalent to local o-minimality in the case of ordered fields.

0.21, p 12, all open sets of $\mathbb{R}$ are equal to an arbitrary disjoint union of open intervals. That is, $O_x \setminus D_i = \bigsqcup_{k \in K} J_k$. (Note that K might be empty.) Now either (a) x is in some J_k or (b) it is not. Consider subcase (a) and let $I := J_k$. Then we know that $I \setminus D_i = I$, which is (iv) above.

In subcase (b), we already have that O_x is an interval around x. But choosing an appropriate I around x that satisfies any of (i)-(iv) is more difficult than in subcase (a). For, it seems possible that all subneighborhoods U_x of O_x, contain infinitely many J_k's. For example, if $x = 0$ and $D_i = [1, 1/2] \cup [1/4, 1/8] \cup [1/16, 1/32] \cup ... \cup \{0\}$, then any subneighborhood U_x will contain infnitely many J_k's. But, in fact, such cases are not possible. More precisely, I claim that $O_x \setminus D_i$ is a *finite* union K_0 of open intervals, $\bigsqcup_{k \in K_0} J_k$. For, suppose not. Then there are infinitely many J_k's in O_x. We show that if $\mathcal{R}$ contains such a definable set D_i, it must contain another definable set D_j such that $\mathcal{R}$ does not model S under V_j at some point y. (This would show that any structure containing D_i must violate the basic assumption for the right-to-left direction that $\mathcal{R}$ models S under *all* definable valuations at every point.) In particular, we let $D_j := D_i \setminus \{x\}$. (We know that this set is definable because we have parameters.) We let V_j be the definable valuation such that $V_j(p) = D_j$. I claim that $V_j, \mathcal{R}, x \not\models \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$.

For contradiction, suppose first that $V_j, \mathcal{R}, x \models \Box(p \vee \Box\neg p)$. It follows that $V_j, \mathcal{R}, x \models p \vee \Box\neg p$. Since $x \notin D_j$, it follows that $V_j, \mathcal{R}, x \models \Box\neg p$. So there is an open $O_x \ni x$ such that $O_x \setminus D_j = O_x$. But since $D_i = D_j \sqcup \{x\}$, we then have that $O_x \cap D_i = \{x\}$. Hence, there are two closest open intervals J_a, J_b, such that $O_x = J_a \sqcup \{x\} \sqcup J_b$[67], which contradicts our assumption that there are infinitely many J_k's in O_x. In this case there are only two: J_a and J_b. Hence, by reductio, we have $V_j, \mathcal{R}, x \not\models \Box(p \vee \Box\neg p)$. Now suppose that $V_j, \mathcal{R}, x \models \Box(\neg p \vee \Box p)$. Then there is an O_x such that for all $y \in O_x$, $V_j, \mathcal{R}, y \models \neg p \vee \Box p$. The second disjunct tells us that $D_j \cap O_x$ must be an arbitrary union of disjoint open intervals $\bigsqcup_{m \in M} L_m$. But we already have that $O_x \setminus D_i$ is an infinite union of disjoint open intervals $\bigsqcup_{t \in T} J_t$. Since $D_i = D_j \sqcup \{x\}$, we then have that $O_x \setminus D_j = (\bigsqcup_{t \in T} J_t) \sqcup (\{x\})$. Hence $O_x = (\bigsqcup_{t \in T} J_t) \sqcup (\bigsqcup_{m \in M} L_m) \sqcup (\{x\})$. And since $|T| \geq \omega$, this is impossible. (To see this, consider the set $O_x \setminus \{x\}$. It can be written as two disjoint open intervals $]a, b[,]b, c[$. Since the first interval is connected, it cannot be the disjoint union of open sets, so it must be identical to either a J_t or an L_m. This means the second, connected interval is equal to an infinite disjoint union of open sets, which is a contradiction.) Hence, $V_j, \mathcal{R}, x \not\models \Box(p \vee \Box\neg p) \vee \Box(\neg p \vee \Box p)$. This violates our supposition for the left-to-right direction. So we can conclude that $O_x \setminus D_i$ is a finite union K_0 of open intervals, $\bigsqcup_{k \in K_0} J_k$.

Recall from the set up of subcase (b), we have that $x \in D_i, x \in O_x$, and $x \notin J_k$ for all k. Since there are finitely many J_k's we can select a open interval $U_x \subseteq O_x$ that contains at most one J_k on either side of x. There are four cases to consider: (1) U_x contains one interval J_a above x and one interval J_b below x, (2) U_x contains one interval J_a above x and is disjoint from all J_k's below x, (3) U_x contains one interval J_b below x and is disjoint from all J_k's above x, and (4) U_x is disjoint from all J_k's.

[67] So everyone is clear, I am using the square cup to mean "disjoint union".

In (4), we can immediately conclude that $U_x \cap D_i = U_x$, which means that U_x satifies (iv). In (3), either $x \in cl(J_b)$ or not. If it is in the closure, then $U_x \cap D_i = [x, d[$ for some d, which means U_x satisfies (iii). If not, then there is an open interval I around x that is disjoint from all J_k's, which means I satisfies (iv). In (2), we have the same as (3). In (1), either $x \in cl(J_a \cup J_b)$ or not. If not, then there is an open I around x which satisfies (iv). If so, then either $x \in cl(J_a) \cap cl(J_b)$ or not. If not, then we can find an I that satisfies (ii) or (iii). If so, then U_x satisfies (i).

In case (II), we follow a similar argument as in case (I). We assume that $V_i, \mathcal{R}, x \models \Box(\neg p \vee \Box p)$. Let $D_k := \mathbb{R} \setminus D_i$. Then we know that $V_k, \mathcal{R}, x \models \Box(p \vee \Box \neg p)$. Since we let V_i be arbitrary in case (I), we can substitute V_k for V_i in the above reasoning. Thus, we have that at x, there is an I such that either $I \setminus D_k$ of $I \cap D_k$ is equal to one of (i)-(iv) above. Since $I \setminus D_k = I \cap D_i$ and $I \cap D_k = I \setminus D_i$, we can say the same of D_i itself.

Using the above-mentioned results from Toffalori and Vozoris (2009) and Schoutens (2014), we can conclude that $\mathcal{R}$ is o-minimal. $\square$

3.8 Bibliography

Baddeley, A. D., Hitch, G.J., and Allen, R.J. (2009). "Working memory and binding in sentence recall." *Journal of Memory and Language* 61, pp 438-456.

Baddeley, A. D., Vallar, G., and Wilson, B. A. (1987). "Sentence comprehension and phonological memory: some neuropsychological evidence." In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading.* London: Erlbaum, pp 509-529.

van Benthem, J. and Bezhanishvili, G. (2007). "Modal Logics of Space." In M. Aiello et al (Eds.), *Handbook of Spatial Logics.* Dordrecht: Springer, pp 217-298.

Bettencourt, K. C. and Xu, Y. (2016). "Decoding the content of visual short-term memory under distraction in occipital and parietal areas." *Nature Neuroscience* 19(1), pp 150-157.

Bolzano, B. (1817). "Purely analytic proof of the theorem that between any two values which give results of opposite sign there lies at least one real root of the equation." In W. Ewald (Ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics, Vol 1.* Oxford: Clarendon, 1996. pp 225-248.

Bower, G. H. and Glass, A. L. (1976). "Structural units and the reintegrative power of picture fragments." *Journal of Experimental Psychology: Human Learning and Memory* 2, pp 456-466.

Brown, J. R. (2008). *Philosophy of Mathematics.* Second Edition. New York: Routledge.

Buechler, S. (2017). *Essential Stability Theory.* Cambridge: Cambridge University Press, 2002.

Burge, T. (1993). "Content Preservation." *The Philosophical Review* 102(4). Reprinted in *Philosophical Issues*, Vol 6, 1995. pp 271-300.

Burge, T. (1998). "Computer Proof, Apriori Knowledge, and Other Minds." *Philosophical Perspectives*, Vol. 12, pp 1-37.

Burge, T. (2003). "Perceptual Entitlement." *Philosophy and Phenomenological Research* 67(3), pp 503-548.

Christophel, T. B., Cichy, R. M., Hebart, M. N., and Haynes, J. (2015). "Parietal and early visual cortices encode working memory content across mental transformations." *NeuroImage* 106, pp 198206.

Coffa, A. (1982). "Kant, Bolzano, and the Emergence of Logicism." *Journal of Philosophy* 79(11): pp 679-689.

Cooper, L.A. and Shepard, R.N. (1973). "Chronometric Studies of the Rotation of Mental Images." In W.G. Chase (Ed.), *Visual Information Processing.* New York: Academic Press.

Corballis, M. G. and Beale, I. L. (1976). *The psychology of left and right.* Hillsdale, N.J.: Erlbaum.

Corballis, M. G. and Beale, I. L. (1983). *The ambivalent mind: The neuropsychology of left and right.* Chicago: Nelson-Hall.

Croom, F. H. (1978). *Basic Concepts of Algebraic Topology.* New York: Springer-Verlag, 2002.

De Toffoli, S. and Giardino, V. (2014). "Forms and Roles of Diagrams in Knot Theory." *Erkenntnis* 79, pp 829-842.

Dretske, F. (2000). "Entitlement: Epistemic Rights without Epistemic Duties?" *Philosophy and Phenomenological Research* 60(3), pp 591-606.

van den Dries, L. (1984). "Remarks on Tarski's problem concerning (R, +, *, exp)." In Lolli, Longo, and Marcja (Eds.), *Logic Colloquium '82.* North-Holland: Elsevier, 1984.

van den Dries, L. (1998). *Tame Topology and O-minimal Structures.* London Mathematical Society Lecture Note Series 248. Cambridge: CUP, 1998.

van den Dries, L. and Speissegger, P. (2003). "O-minimal Preparation Theorems." Proceedings of the Euro-Conference in Model Theory and Applications, Ravello, Italy.

Finke, R. A., Pinker, S., and Farah, M. J. (1989). "Reinterpreting Visual Patterns in Mental Imagery." *Cognitive Science* 13, pp 51-78.

Folland, G. (1999). *Real Analysis: Modern Techniques and Their Applications.* New York: Wiley & Sons, 1999.

Forder, H. (1927). *The Foundations of Euclidean Geometry.* Cambridge: Cambridge University Press.

Fraenkel, A. (1927). *Zehn Vorlesungen ber die Grundlegung der Mengenlehre.* Leipzig: B. G. Teubner.

Giaquinto, M. (1994). "Epistemology of Visual Thinking in Elementary Real Analysis." *The British Journal for the Philosophy of Science*, Vol. 45, No. 3, pp 789-813.

Giaquinto, M. (2005a). "From Symmetry Perception to Basic Geometry." In P. Mancosu (Ed.), *Visualization, Explanation, and Reasoning.* Oxford: Oxford University Press, pp 31-56.

Giaquinto, M. (2005b). "Mathematical Activity." In P. Mancosu (Ed.), *Visualization, Explanation, and Reasoning Styles in Mathematics.* Dordrecht: Springer, pp 75-89.

Giaquinto, M. (2007). *Visual Thinking in Mathematics*. Oxford: Oxford University Press, 2007.

Giaquinto, M. (2008a). "Visualizing in Mathematics." In P. Mancosu (Ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp 22-42.

Giaquinto, M. (2008b). "Cognition of Structure." In P. Mancosu (Ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp 43-64.

Giaquinto, M. (2011). "Crossing Curves: A Limit to the Use of Diagrams in Proofs." *Philosophia Mathematica* 19(3), pp 281-307.

Gödel, K. (1964). "What is Cantor's continuum problem?" In S. Feferman et al (Eds.), *Collected Works Vol II*. Oxford: Oxford University Press, 1990. pp 254-270.

Goldman, A. (2001). "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63(1), pp 85-110.

Gray, J. J. (2009). "Nineteenth Century Analysis as Philosophy of Mathematics." In B. van Kerkhove (Ed.), *New Perspectives on Mathematical Practice*. Singapore: World Scientific, pp 138-149.

Grothendieck, A. (1983). "Pursuing Stacks". http://www.math.umn.edu/∼messing/docs/Grothendieck%20–%20Pursuing%20stacks.pdf

Grothendieck, A. (1984). "A Sketch of a Programme." In L. Schneps and P. Lochak (Eds.), *Geometric Galois Actions I: Around Grothendieck's Esquisse D'un Programme*, London Mathematical Society Lecture Note Series 242, Cambridge University Press, 1997.

Hahn, H. (1933). "The Crisis in Intuition." In J. R. Newman (Ed.), *The World of Mathematics, Vol 3.* New York: Simon and Schuster, 1956. pp 1956-1976.

Hamamé, C. M., Vidal, J. R., et al. (2012). "Reading the mind's eye: Online detection of visuo-spatial working memory and visual imagery in the inferior temporal lobe." *NeuroImage* 59, pp 872-879.

Harris, I. M. and Miniussi, C. (2003). "Parietal Lobe Contribution to Mental Rotation Demonstrated with rTMS." *Journal of Cognitive Neuroscience* 15 (3), pp 315-323.

Haskell, D. and Pillay, A. and Steinhorn, C. (2000). "Overview." In: *Model Theory, Algebra, and Geometry.* Mathematics Sciences Research Institute. Cambridge: CUP, 2000.

Heis, J. (2008). *The Fact of Modern Mathematics: Geometry, Logic, and Concept Formation in Kant and Cassirer.* Doctoral Dissertation, University of Pittsburgh.

Hochberg, J. and Gellman, L. (1977). "The effect of landmark features on 'mental rotation' times." *Memory and Cognition* 5: pp 23-26.

Jolicoeur, P. and Kosslyn, S. M. (1983). "Coordinate systems in the long-term memory representation of three-dimensional shapes." *Cognitive Psychology* 15, pp 301-345.

Jordan, K., Heinze, H.-J., Lutz, K., Kanowski, M., and Jäncke, L. (2001). "Cortical Activations during the Mental Rotation of Different Visual Objects." *NeuroImage* 13, pp 143-152.

Klein, F. (1884). *Lectures on the Icosahedron.* New York: Dover, 2003.

von Koch, H. (1906). "Une méthode géométrique élémentaire pour l'étude de certaines questions de la théorie des courbes planes." *Acta Mathematica* 30: pp 145-174.

Kosslyn, S. M. and Shwartz, S. P. (1977). "A Simulation of Visual Imagery." *Cognitive Science* 1, pp 265-295.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate.* Cambridge: MIT Press.

Kreisel, G. (1967). "Informal Rigour and Completeness Proofs." In Lakatos (Ed.), *Problems in the Philosophy of Mathematics.* Proceedings of the International Colloquium in the Philosophy of Science, London, 1965. North-Holland Publishing, pp 138-171.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery.* Cambridge: Cambridge University Press.

Logie, R. H. (1995). *Visuo-spatial working memory.* Hove, UK: Lawrence Erlbaum.

Logie, R. H. and Cowan, N. (2015). "Perspectives on working memory." *Memory and Cognition* 43, pp 315-324.

Logie, R. H. and van der Meulen, M. (2009). "Fragmenting and integrating visuospatial working memory." In Brockmole, J. R. (Ed.), *The Visual World in Memory.* Hove, UK: Psychology Press, pp 1-32.

Lützen, J. (2003). "The Foundation of Analysis in the 19th Century." In *A History of Analysis*, volume 24 of History of Mathematics. American Mathematical Society, Providence, RI, 2003. pp 155-195.

Mancosu, P. (2005). “Visualization in Logic and Mathematics.” In P. Mancosu (Ed.), *Visualization, Explanation, and Reasoning Styles in Mathematics.* Dordrecht: Springer, pp 13-30.

Manders, K. (1996). “Diagram Contents and Representational Granularity.” In J. Seligmann and D. Westerstahl (Ed.), *Logic, Language, and Computation.* Stanford: CSLI, pp 389-404.

Manders, K. (2008a). “The Euclidean Diagram.” In P. Mancosu (Ed.), *The Philosophy of Mathematical Practice.* Oxford: OUP, pp 80-133.

Manders, K. (2008b). “Diagram-Based Geometric Practice.” In P. Mancosu (Ed.), *The Philosophy of Mathematical Practice.* Oxford: OUP, pp 65-79.

Marker, D. (2002a). “Review of *Tame Topology and O-minimal Structures*”. Bulletin of the American Mathematical Society. Volume 27, Number 3, pp 351-357.

Marker, D. (2002b). *Model Theory: An Introduction.* New York: Springer-Verlag, 2002.

McMullen, P. A. and Farah, M. J. (1991). “Viewer-centered and object-centered representations in the recognition of naturalistic line drawings.” *Psychological Science* 2, pp 275-277.

Munkres, J. R. *Topology.* Second Edition. Upper Saddle River: Prentice Hall Inc., 2000.

Norman, J. (2006). *After Euclid: Visual Reasoning and the Epistemology of Diagrams.* Stanford: CSLI Publications.

Palmer, S. E. (1977). “Hierarchical Structure in Perceptual Representation.” *Cognitive Psychology* 9, pp 441-474.

Palmer, S. E. (1980). “What Makes Triangles Point: Local and Global Effects in Configurations of Ambiguous Triangles.” *Cognitive Psychology* 12, pp 285-305.

Palmer, S. E. (1983). “The psychology of perceptual organization: A transformational approach.” In J. Beck et al (Ed.), *Human and Machine Vision.* New York: Academic Press, 1983.

Palmer, S. E. (1985). “Palmer, S. (1985). The role of symmetry in shape perception.” *Acta Psychologica* 59, pp 67-90.

Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology.* Cambridge: MIT Press, 1999.

Palmer, S. E. and Hemenway, K. (1978). “Orientation and symmetry: Effects of multiple, rotational, and near symmetries.” *J Exp Psychol Hum Percept Perform* 4(4), pp 691-702.

Peacocke, C. (1992). *A Study of Concepts*. Cambridge: MIT Press, 1992.

Pearson, D. G. (2001). "Imagery and the visuo-spatial sketchpad." In J. Andrade (Ed.), *Working memory in perspective*. Hove, UK: Psychology Press, pp 33-59.

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S., Burnett Heyes, S., Holmes, E. A. (2013). "Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework." *Clinical Psychology Review* 33, pp 1-23.

Peterzil, Y. and Starchenko, S. (1998). "A trichotomy theorem for o-minimal structure." *Proc. London Math. Soc.* 77(3), pp 481-523.

Peterzil, Y. (2007). "A Self-Guide to O-minimality." Camerino Tutorial June 2007.

Plato. *Meno*. In J. Cooper (Ed.), *Plato: Complete Works*. Hackett, 1997.

Poincaré, H. (1895). "Analysis Situs." *Papers on Topology: Analysis Situs and Its Five Supplements*. Translator: John Stillwell. USA: American Mathematical Society, 2010. pp 5-74.

Prinz, J. J. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis.* Cambridge: MIT Press.

Rennett, A. (2014). "The non-axiomatizability of o-minimality." *The Journal of Symbolic Logic*, 79(1), pp 54-49.

Rosch, E. and Mervis, C. B. (1975). "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7, pp 573-605.

Rudin, W. (1976). *Principles of Mathematical Analysis.* Third Edition. New York: McGraw Hill.

Russell, B. (1901). "Mathematics and the Metaphysicians." In: B. Russell, *Mysticism and Logic*. London: George Allen and Unwin Ltd, 1910. pp 74-96.

Russell, B. (1919). *Introduction to Mathematical Philosophy.* London: George Allen and Unwin; New York: The Macmillan Company.

Schoutens, H. (2014). "O-minimalism." *Journal of Symbolic Logic* 79(2), pp 355-409.

Shepard, R.N. and Metzler, J. (1971). "Mental Rotation of Three-Dimensional Objects." *Science* (171), pp 701-703.

Sinaceur, H. (1994). "Calculation, Order, and Continuity." In P. Ehrlich (Ed.), *Real Numbers, Generalizations of the Reals, and Theories of Continua.* Netherlands: Kluwer, pp 191-206.

Steinhorn, C. (2003). "A Brief Introduction to O-minimality." In Edmundo (ed). *O-minimal Structures, Lisbon 2003: Proceedings of a Summer School by the European Research and Training Network RAAG.* pp 11-31.

Tarski, A. (1938). "Sentential Calculus and Topology". Printed in: *Logic, Semantics, and Metaphysics.* Oxford: Clarendon Press, 1983.

Tarski, A. (1948). "A Decision Method for Elementary Algebra and Geometry." The RAND Corporation. Second Edition, 1957.

Tarski, A. (1959). "What is Elementary Geometry?" In L. Henkin et al (Eds.), *The axiomatic method.* Amsterdam: North-Holland, pp 16-29.

Toffalori, C. and Vozoris, K. (2009). "Notes on local o-minimality." *Mathematical Logic Quarterly* 55(6), pp 617 - 632.

Tye, M. (1991). *The Imagery Debate.* Cambridge: MIT Press.

Wilkie, A. J. (1996). "Model Completeness Results for Expansions of the Ordered Field of Real Numbers by Restricted Pfaffian Functions and the exponential function." *J American Math Soc* 9(4), pp 1051-1094.

Wilkie, A.J. (2007). "O-minimal Structures." *Seminarie BOURBAKI.* pages 985-01 to 985-11.

Yomdin, Y. and Comte, Ge. (2004). *Tame Geometry with Application in Smooth Analysis.* Lecture Notes in Mathematics (1834). New York: Springer.

Zacks, J. (2007). "Neuroimaging Studies of Mental Rotation: A Meta-analysis and Review." *Journal of Cognitive Neuroscience* 20(1), pp 1-19.