

# A Preliminary Model of the Development of Early Visual Perception

Kristofer Brooking

A thesis submitted to

Auckland University of Technology

in partial fulfilment of the requirements for the degree of  
Master of Computing and Information Sciences (MCIS)

2015

School of Computer and Mathematical Sciences

## **ABSTRACT**

Developmental researchers typically use equivocal terminology to explain the foundation of visual perception in infancy. Terms such as “occlusion”, “permanence”, and even “object” provide a convenient explanation of infant behaviour to a learned audience. However, they also afford the infant unstated, and perhaps undeserved knowledge of the world. Are infants born with the ability to individuate and identify mid-sized objects? How is an object distinguished from its surroundings? This thesis seeks to explore potential explanations for infant behaviours without using ambiguously leading terminology, and in so doing, understand the explanatory power of these post-hoc linguistic definitions of object knowledge.

This thesis investigates some of these questions through an interpretive approach that is applied to behavioural experiments involving the development of visual perception in infancy. These experiments typically consist of a procedure in which object knowledge is attributed to infants based on preferential looking towards one of two different displays. What can be obscured or ignored by the interpretations of developmental researchers is the possibility that object knowledge has developed from lower level primitives. This thesis attempts to analyse object-related concepts and describe them “computationally” using primitives such as position and trajectory.

A bottom-up approach to interpreting visual perception is developed where a concept is described by its observed input and output variables, and a plausible mechanism for the conversion from the former to the latter is deduced. Sensory input consists of low level positional change and simple motor action. Higher level mechanisms are then built on top of this foundation. The bottom-up interpretive approach was developed iteratively during the course of this research, and is considered a novel contribution of the thesis.

The bottom-up interpretation of multiple experiments over several iterations of the research process resulted in a preliminary theoretical model of early visual perception. The model consists of perceptual primitives that explain commonly assumed object-related principles including continuity, cohesion, and segregation. Importantly, it does so

using a foundation consisting only of visible positional change and motor action so that a post-hoc linguistic definition can provide no additional explanatory power.

## TABLE OF CONTENTS

ABSTRACT .....	ii
LIST OF FIGURES .....	vii
LIST OF TABLES .....	viii
ATTESTATION OF AUTHORSHIP .....	ix
ACKNOWLEDGEMENTS .....	x
1 INTRODUCTION .....	11
1.1 Background .....	11
1.2 Motivation .....	13
1.3 Research Objectives .....	14
1.4 Novelty and Contribution .....	17
1.5 Thesis Structure .....	17
2 LITERATURE REVIEW .....	19
2.1 Introduction .....	19
2.2 Computational Models of Infant Development .....	20
2.3 Assumed Object Knowledge .....	22
2.3.1 What is an Object? .....	22
2.3.2 The Assumption of Object Persistence .....	24
2.4 The Relation to Mental Representation .....	25
2.4.1 A Brief History of Developmental Research .....	26
2.4.2 Representational Emergence .....	28
2.4.3 The Conflation of Representation and Object Knowledge .....	28
2.5 Methodologically Related Literature .....	32
2.5.1 Re-Interpretation of the Drawbridge Experiment .....	32
2.5.2 Re-Interpretation of the Short and Tall Object Experiment .....	34
2.6 Computational and Algorithmic Analysis .....	36
2.7 Summary .....	38
3 RESEARCH DESIGN AND IMPLEMENTATION .....	40
3.1 Introduction .....	40
3.2 Exploratory Research Approach .....	41
3.3 Research Methods .....	43
3.4 Methodology, Theoretical Perspective, and Epistemology .....	44

3.4.1	Methodology for Phase 1 – Data Search and Initial Analysis .....	44
3.4.2	Methodology for Phase 2 – Re-Interpretations of Experiments.....	45
3.4.3	Theoretical Perspective and Epistemology .....	46
3.5	Research Implementation .....	47
3.5.1	Data Search Protocol.....	47
3.5.2	Data Analysis .....	48
3.5.2.1	Data Reduction .....	49
3.5.2.2	Data Display and Comparison .....	49
3.5.2.3	Conclusion Drawing .....	50
3.5.2.4	Experiment Re-Interpretation Protocol.....	50
3.5.3	Iterative Updates to the Methodology.....	52
3.5.4	Limitations .....	52
4	FINDINGS AND DISCUSSION.....	54
4.1	Introduction .....	54
4.2	Experiment 1 – Spatiotemporal Completion (Eye Tracking).....	58
4.2.1	Summary of Original Experimental Procedure.....	58
4.2.2	Results and Original Interpretation .....	59
4.2.3	Naïve Interpretation .....	59
4.2.4	Discussion and Justification.....	62
4.3	Experiment 2 - Spatial Completion (Including Eye Tracking) .....	64
4.3.1	Summary of Original Experimental Procedure.....	64
4.3.2	Results and Original Interpretation .....	64
4.3.3	Naïve Interpretation .....	66
4.3.4	Discussion and Justification.....	68
4.4	Experiment 3 – Spatiotemporal Completion .....	71
4.4.1	Summary of Original Experimental Procedure.....	71
4.4.2	Results and Original Interpretation .....	72
4.4.3	Naïve Interpretation .....	72
4.4.4	Discussion and Justification.....	78
4.5	Experiment 4 – Object Individuation and Identity.....	80
4.5.1	Summary of Original Experimental Procedure.....	80
4.5.2	Results and Original Interpretation .....	81

4.5.3	Naïve Interpretation .....	82
4.5.4	Discussion and Justification.....	84
4.6	General Discussion of the Naïve Model.....	85
4.6.1	Summary of the Model.....	85
4.6.2	Perceptual Primitives .....	88
4.6.3	Bottom-Up Object Knowledge .....	90
4.6.4	Developing the Model Further .....	92
4.6.4.1	Experiment 5 – Object Permanence.....	92
4.6.4.2	Results and Original Interpretation.....	94
4.6.4.3	Discussion.....	94
5	CONCLUSION.....	99
5.1	Summary .....	99
5.2	The Bottom-Up Interpretive Approach .....	100
5.3	The Naïve Model.....	101
5.4	Novelty and Contribution.....	104
5.5	Limitations.....	104
5.6	Recommendations .....	105
5.7	Future Work .....	105
6	REFERENCES.....	106
7	GLOSSARY.....	111

## LIST OF FIGURES

Figure 1: Drawbridge Experiment .....	32
Figure 2: Short and Tall Object Experiment .....	34
Figure 3: Spatiotemporal Completion (Eye Tracking).....	58
Figure 4: Spatial Completion (Including Eye Tracking).....	64
Figure 5: Spatiotemporal Completion.....	71
Figure 6: Occlusion Event.....	74
Figure 7: Segregation Mechanism .....	75
Figure 8: Object Individuation Familiarisation Events.....	80
Figure 9: Object Individuation Test Events .....	80
Figure 10: Object Permanence Habituation Events .....	92
Figure 11: Object Permanence Impossible Test Event .....	93
Figure 12: Object Permanence Possible Test Event .....	93

## LIST OF TABLES

Table 1: Data Search Protocol.....	48
Table 2: Summary of original interpretations of results for experiment 1.....	59
Table 3: Summary of original interpretations of results for experiment 2.....	65
Table 4: Summary of original interpretations of results for experiment 2(eye tracking).....	66
Table 5: Summary of original interpretations of results for experiment 3.....	72
Table 6: Summary of original interpretations of results for experiment 4.....	82
Table 7: Summary of original interpretations of results for experiment 5.....	94

## **ATTESTATION OF AUTHORSHIP**

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

## **ACKNOWLEDGEMENTS**

I wish to acknowledge and thank my supervisor, Prof. Wai Yeap, for his continued support over the past year. His guidance has been invaluable, and is greatly appreciated.

# 1 INTRODUCTION

## 1.1 Background

We take visual perception for granted and rarely stop to consider how something so remarkable works so effortlessly. We open our eyes and the light arriving at our retinas is interpreted as what we understand to be reality. It is this understanding that is remarkable considering that the retinas themselves are analogous to cameras, and cameras clearly cannot infer meaning from a series of images. The human brain somehow takes the stream of images and translates them into our perception of the world; raw sensory information is processed and a meaningful representation of that information is presented to the observer. This mental representation is about the physical world, but it is not necessarily exactly the same thing.

Intentionality is the fundamental philosophical issue of “aboutness” (Siewert, 2002). The term refers to the rather ambiguous notion that something is about or directed toward something else, and the philosophical issue is how a thing comes to be about another thing. It is often debated in the context of mental representation (Siewert, 2002); if one thinks of an object, it exists differently in the mind than it does in reality.

Depending on the theory there can also be different types of intentionality. Things that have original intentionality (like thoughts) are inherently about something, whereas things with derived intentionality are about something only because something or someone with original intentionality provided it (Byrne, 2006). The puzzle is how something comes to have original intentionality.

In Brentano’s (1874) view, only mental phenomena can exhibit intentionality because they are by definition about or directed toward something else. “Intentional inexistence” refers to objects or mental states existing only in the mind. Further to this, Brentano posits that all mental acts are intentional and that the intentional object of a mental state need not be the same as that which it is about (Brentano, 1874). Essentially, he is proposing that our view of the world is not veridical, instead it is a subjective account based on our thoughts and ideas.

As an example, right now I am looking at my computer monitor. The intentional object that I am representing is that of a widescreen monitor plugged into my laptop that gives me more room to work. Even as I type this with the entire monitor in my line of sight, I could not say what the model number is (it's displayed on the top right corner) or the number of configuration buttons that are visible along the bottom edge. I certainly don't know the weight of the monitor or the number of plastic panels that are used in its construction. My representation of the monitor is not complete. I'm only provided with the information that I require which currently happens to be my work on this document.

Brentano (1874) states that intentional objects do not have to be associated with any physical thing. Indeed, this is one of the distinctions between physical and mental phenomena. I can think about a thing without that exact thing existing in the real world. If I close my eyes for a while or look away from my desk, the computer does not cease to exist. The physical thing remains as it was and the intentional object exists in my mind. The representation might have changed significantly, but I am still aware that there is a monitor on my desk.

Brentano may have actually been referring to the idea that a thought can just as easily be about a computer as about a unicorn, but this thesis is more concerned with visual perception than with imagination. The important idea for either scenario is that the intentional object need not be exactly the same as its real-life equivalent.

If we subscribe to this theory of intentional perception, the next logical scientific step is to ask how it might operate. To determine how physical objects are represented as a mental state, the scientific method would suggest that reduction may be useful. This is the crux of the intentionality debate. It is not, and may not ever be known whether "aboutness" can be explained by way of a reduction to physical brain activity. Even if intentionality is reducible, brain activity is interpreted at the lowest electrical and chemical levels. Attempting to gain insight into intentionality from this alone is akin to trying to decipher the messaging of a computer network given only the flickering lights on the front of a switching hub.

There is another type of empirical data that is useful for the interpretation of mental representations. It is observed by way of externally manifested responses to visual stimuli.

Psychologists use behavioural experiments to examine many areas of visual perception. However, merely observing the external response to a stimuli in an already fully functional system cannot by itself give much information as to the development of the system. More insight can be gained through testing different stages of the development of the system and then comparing the results. Fortunately, developmental researchers have been performing behavioural experiments involving young infants for decades and there is abundant literature on many different aspects of visual perception. Evidence from these experiments fuels one of the fundamental debates in psychology: what is most important to the development of intelligence? Nature or nurture. Is crucial worldly understanding built into the human schematic or is learning alone able to explain the acquisition of knowledge from newborn baby to adult?

## **1.2 Motivation**

Developmental research has intersected with the fields of computer science and robotics (Lungarella, Metta, Pfeifer & Sandini, 2003). Until recently the majority of computer science research into intelligence has focused on solving problems generally associated with a single domain of knowledge. The purpose of machine learning is typically to classify and forecast or predict given large amounts of similar data. The models built for this purpose are typically pre-programmed or trained offline in order to complete the problem that they were designed to solve. This pre-programming or training is typically not generalisable to other different, or in many cases even similar problems in other domains.

The field of developmental artificial intelligence has been gaining momentum over the past few decades as a potential solution to this problem (Guerin, 2011). The idea is that if a computational model is given only the algorithms necessary to learn domain general information, then it could develop similarly to an infant and pick up common sense knowledge by itself (Guerin, 2011). This common sense knowledge is fundamental to the generalisation of understanding from one domain to another. It is also an important aspect of “Strong AI” in which a machine would be able to match a human intellectually given any cognitive task.

This thesis is ultimately motivated by an interest in furthering the understanding of the development of general intelligence. Visual perception was selected as the focus of this research because it seems to be one of the foundational systems of human intelligence, and can be explored at a low level. The visual world provides meaning to the infant long before the linguistic ability to describe it has begun to develop. It is plausible that the interactions of objects in space are a precursor to higher level linguistic knowledge (Mandler, 2012). Visual perception is also a good platform to think about intentionality. What does it mean to perceive an object and encode it into a mental representation?

### **1.3 Research Objectives**

This thesis began with the exploratory objective of finding the fundamental issues with the computational approaches to simulating visual development in infancy. It became clear that an issue inherent to many of these computational models is that they make assumptions about object knowledge based on the approach developmental researchers take to analysing cognition. Computational models often seem to be built to agree with developmental theories rather than to question them (Roberts & Pashler, 2000). Therefore, the first research objective is to identify the assumptions of infant knowledge and the potential effects they have on our understanding of visual perception.

The developmental literature typically attempts to explain intelligence at a high level of cognition, and while useful to the analysis of human behaviour, this is of limited use in explaining low level domain general systems of learning. Developmental researchers use terms like “object permanence” with the implication that the understanding of an infant is the same as that of an adult. This cannot be known because clearly young infants are unable to articulate a response to experiments. Infants do show predictable responses to stimuli, but there is no way to know the cause. What is often obscured or ignored by developmental researchers is that high level cognition must have at some point developed from lower level primitives. This thesis aims to explore potential explanations for infant behaviours without using ambiguously leading terminology.

Developmental researchers typically use terms such as occlusion, permanence, and motion to explain the fundamental concepts that form the foundation of visual perception. If the empirical evidence from behavioural experiments can be re-interpreted using

perceptual primitives below this level of understanding, new insight into perception may emerge. This thesis attempts to analyse these concepts and describe them “computationally” using perceptual primitives such as position and trajectory. A bottom-up approach to interpreting cognition is employed where sensory information is the starting point and higher level abilities are built on top of this foundation.

As a quick example, consider the following experiment. An infant is seated in front of an empty stage. A researcher reaches in from behind a curtain and places a doll on the stage. A screen slides down in front of the doll so that the infant can no longer see it. Another hand comes in from the side of the stage holding another similar doll, and places it behind the screen. The screen is removed and reveals either one doll (an impossible event), or two dolls (a possible event) sitting on the stage. Infants are interpreted as maintaining a representation of the dolls while hidden from view if more time is spent looking at the impossible than the possible event. The interpretation is that looking preferentially at the impossible event means that the infant is surprised that one of the dolls has disappeared from existence. With this experimental procedure five-month-old infants were interpreted as maintaining a representation of these objects while occluded from view (Wynn, 1992).

On the surface this seems like clear empirical evidence that five-month-old infants are able to maintain a mental record of the number of dolls placed on the stage. However, this requires attributing the infant with a fair amount of conceptual knowledge. Firstly, the infant needs to know that dolls don’t disappear from existence when out of sight. To an adult, this is an unstated fact and doesn’t require conscious thought, but why would an infant automatically know this? Next, the infant needs to maintain a representation for both dolls individually. This doesn’t necessarily have to mean that the infant pictures two dolls behind the screen, but some type of placeholder is required. Finally, when the screen reveals the stage, the infant needs to be able to re-identify the doll or dolls as the same objects that are in the mental representation. Otherwise there cannot possibly be a reaction of surprise to an unexpected number of dolls; either both options are unexpected or neither are.

Further questions arise if we dig a little deeper. How is an occlusion event perceived by an infant? How does the infant learn to differentiate occlusion from disappearance? Is the

doll on the stage perceived as separate from the background? Does this matter? Clearly, most of these questions are unimportant to an average adult perceiver who understands object knowledge instinctively. Yet these questions may have a big impact on the progression of development through infancy.

The depth to these questions makes designing a research project rather difficult. This is an interpretive and exploratory project so there is no clear quantifiable metric for success or failure. The scope needs to be open enough to allow useful findings to emerge; however, constraints on scope are necessary so that the research follows a coherent direction. Therefore, this research will only attempt to look into a single object-related concept. This both narrows the scope and leaves room to explore any perceptual processes that are the foundation for this concept. At a high level the research is intended to be about intentionality or mental representation. From a developmental research perspective this aligns closely with experiments that test object permanence – a conceptual understanding that a hidden object continues to exist in the world after its apparent disappearance.

As is typically the case with exploratory research, a tentative research question is provided but is secondary to more general research objectives. This provides additional opportunity for discovery at the expense of a less structured research process. Consequently, the research objectives require careful forethought so that the original problem doesn't lose focus during the lifetime of the project.

The research objectives are:

- To identify the assumptions of infant knowledge and the potential effects they have on our understanding of visual perception, and use this to build a bottom-up interpretive approach that can be applied to developmental experiments.
- To use the bottom-up interpretive approach to explain the perceptual primitives responsible for the development of early object knowledge.

The tentative research question is:

- Can a model built using perceptual primitives describe the empirical evidence from infant behavioural experiments related to the development of object permanence?

## **1.4 Novelty and Contribution**

A bottom-up approach to interpreting infant behavioural experiments will be developed. The approach is intended to provide insight to an often unstated issue in developmental research – the development of early object knowledge. The bottom-up approach uses the observed input and output variables of a concept to deduce the likely mechanism for the conversion from the former to the latter. Sensory input consists of low level positional change, and simple motor action. Higher level mechanisms are then built on top of this foundation. The bottom-up interpretive approach is developed iteratively during the course of the research, and so is considered a novel contribution of the thesis.

A separate contribution will be the model of development that results from the bottom-up interpretation of several behavioural experiments. More insight into the development of object knowledge will be available if the relationships between multiple interpretations can be found. This will require the careful analysis of multiple developmental studies.

## **1.5 Thesis Structure**

The thesis consists of the following five chapters:

**INTRODUCTION** – Chapter 1 provides the background, motivation, and objectives for the thesis including an example of the unstated assumptions common to interpretations of early object knowledge.

**LITERATURE REVIEW** – Chapter 2 establishes that the majority of developmental research relies on interpretations based on innate object knowledge, and that a bottom-up perceptual approach to interpretation is required to investigate this more thoroughly. Additionally, the review attempts to conceptually ground the important aspects of the research objective so that the terminology is not later misconstrued.

**RESEARCH DESIGN AND IMPLEMENTATION** – Chapter 3 justifies the exploratory research approach and methods required for the research objective; and describes the implementation of the methods in detail.

**FINDINGS AND DISCUSSION** – Chapter 4 presents the findings as a list of the most significant interpretations. It compares the findings with similar research from the literature review.

**CONCLUSION** – Chapter 5 presents the significance of the findings, and identifies the limitations inherent to the exploratory research process.

## 2 LITERATURE REVIEW

### 2.1 Introduction

The literature review will argue that the majority of developmental research is conducted using a theoretical framework that makes fundamental assumptions about infant knowledge, and that these assumptions can be removed with a non-conceptual approach to interpreting visual perception.

Section 2.2 “Computational Models of Infant Development” explains the rationale behind the research objective, and its relation to computational modelling. It is claimed that computational models of developmental theory are configured with innate assumptions about objects. These assumptions are knowingly or unknowingly transferred from the developmental research that they are built to support.

Section 2.3 “Assumed Object Knowledge” discusses the often ambiguous definition for “object” within the developmental literature, and establishes that object principles are not explanatory except as a linguistic description that can be applied after perception has occurred.

Section 2.4 “The Relation to Mental Representation” presents a brief history of developmental research and contends that assumed object knowledge is a direct consequence of assumed innate mental representation. This discussion demonstrates the prevalence of innate representation and therefore of innate object knowledge within the field of developmental research.

Section 2.5 “Methodologically Related Literature” contends that there is limited methodologically similar research because re-interpretations of experiments typically only argue against specific experimental procedures with the objective being to empirically disprove an experiment by repeating it with slight modifications. In contrast, the objective of this research is to theoretically explain the development of the object knowledge involved.

Section 2.6 “Computational and Algorithmic Analysis” describes the foundation of the bottom-up interpretive approach that will be developed. Marr’s computational and

algorithmic levels will be used as the framework. The computational level aligns closely with the identification of assumed object knowledge, while the algorithmic level is used as the basis for the re-interpretation of the events in an experiment.

## **2.2 Computational Models of Infant Development**

Computational models of infant development can aid in understanding psychological theories. Theoretical issues can be tested by creating detailed models based on empirical data from developmental experiments. These models provide the medium for researchers to look at the experiment in more detail by changing learning-related variables that would be difficult or impossible to test with infant subjects (Schlesinger & McMurray, 2012).

Mareschal and Thomas (2007) argue that for a computational model to be a useful tool to a developmental scientist, it must be transparent, grounded and plausible. A model is transparent if the underlying mechanism is well understood; it is grounded if it represents all domains that it is built for equally well; and it is plausible if the underlying mechanisms are consistent with the actual mechanisms that it represents. Mareschal and Thomas note the differences in computational modelling from the perspective of a psychologist and an engineer. An engineer is often more concerned with building a model that succeeds in performing a task than with the plausibility of the model when compared to natural human development. They suggest that the goal is not only to implement a working solution but also to consider how it works and what implications that presents.

Although this is a valid suggestion, it applies differently in the context of the current research. Mareschal and Thomas intend this to apply to the mechanism of learning itself. The issue with many computational models is that they make the same assumptions about object knowledge as the developmental researchers who interpreted the original empirical data (Roberts & Pashler, 2000).<sup>1</sup> Computational models often support the original

---

<sup>1</sup> The objective that was included in the original proposal for this thesis differs from the objective outlined in chapter 1, however, the two are related. The original objective was to review and analyse current computational models of object reasoning, suggest limitations when comparing these to infant development, and offer recommendations for improvement. The belief at the time was that there would be many minor issues that could be discussed. It became clear that a fundamental issue inherent to many computational models is that they rely on the assumption of innate object knowledge. This results in an increased likelihood that the model will support the original interpretations of the empirical data.

interpretations of developmental researchers because they begin at the same cognitive starting point and replicate the same or similar learning mechanisms. Then they use the original developmental evidence as justification for necessary innate object knowledge.

As an example, Mareschal and Johnson (2002) built a connectionist model of Kellman and Spelke's (1983) rod-and-box experiment that aims to learn about object unity – the principle that two disparate sections of an object could be parts of a whole. However, it does so using hard-coded perceptual modules that are akin to innate pattern matchers. The display is split down the middle by two separate motion detectors, allowing a common-motion module to identify exactly when motion occurs in both halves of the display simultaneously. Further modules calculate the axis of each rod part and determined whether an extension of the two rod parts would intersect. Their argument for these features of the model is that infants are equipped with certain object knowledge from birth, and that their objective is not to model the entire process of visual perception. Unfortunately by taking this approach, the model loses some of its explanatory power. Essentially, the model is being built with a specific set of domain knowledge and the objective that it explain the phenomena already assumed to be true.

More recently, Franz and Triesch (2010) designed a unified model for the development of object permanence, object unity, and occluded object trajectory using a modular recurrent neural network. Their model and testing was designed with plausibility in mind, and so included elements like pre-training the network, and prior habituation of the network to the stimulus. The results of their model are supported by evidence from various psychological studies. However, they made the decision to split the foreground and background objects into separate neural populations. As is often the case with psychological theory, computational models of infant development build in mechanisms for object segregation, with the rationale that evidence exists showing figure-ground segregation is available to newborn infants (Slater, Morison, Somers, Mattock, Brown & Taylor, 1990; Slater, Johnson, Brown & Badenoch, 1996). The issue is that it is this original evidence that requires further validation.

## 2.3 Assumed Object Knowledge

### 2.3.1 What is an Object?

An issue with assuming innate object knowledge is that the definition of object is often ambiguous, and differs depending on the developmental research paradigm. In addition, the definition changes during development as different properties of an object are learned. It can also change depending on the task being studied because different object information can be available depending on the situation.

Spelke (1994) proposed that infants make inferences about objects based on three principles: “cohesion (objects move as connected, bounded units), continuity (objects move on connected, unobstructed paths) and contact (objects affect one another’s motion if and only if they touch)”. Mandler (2012) uses different terminology but describes the conceptual primitive **THING** as “any perceptually bounded cohesive object”.

Developmental researchers tend to agree that there are two types of object information: spatiotemporal and featural. Spatiotemporal information consists of the position of the object in space and time. Featural information includes the information you might automatically associate with an object: colour, shape, size, etc. (Wilcox & Baillargeon, 1998; Xu & Carey, 1996). This definition is not a strict separation. For example, the shape of an object can impact its position in space (Krøjgaard, 2004).

In order to identify spatial and featural information, Spelke’s object principles are required. Spatiotemporal information requires object continuity, cohesion, segregation and often re-identification. Without these properties an object in motion might appear as a series of different instances of the same image. In other words, an object cannot be described as an object unless it remains the same instance over time, independent of the rest of the environment.

Without attributing infants with this foundational understanding, it is often difficult to interpret empirical data from experiments. However, the evidence supporting these claims of object knowledge is often inconsistent. In an influential study, Xu and Carey (1996) demonstrated that depending on age, infants use different properties of an object to maintain a representation while occluded. Infants were familiarised with a display in

which two objects, a duck and a ball, would slide out one at a time from behind opposite sides of a screen. The screen would then move to reveal either one or both objects on the stage behind. Ten-month-old infants were shown to ignore a change in featural information, but to react to a change in spatiotemporal information. In contrast, twelve-month-old infants did react to the same change in featural information. Xu and Carey interpreted this as showing that younger infants primarily use spatiotemporal information to maintain representations of occluded objects.

There is contrasting evidence of course. Variances in experimental procedure can dramatically change results. Wilcox and Baillargeon (1998) repeated the experiment using a simpler method in which demands on memory were seemingly reduced. They suggested that infants as young as seven and a half months could individuate an object based on featural information alone. With further changes to experimental procedure, Wilcox (1999) showed that infants aged four and a half months could individuate an object using featural information consisting of size and shape.

The differences in findings based on experimental procedure highlight the difficulty in interpreting infant knowledge from observation alone. Is the experimental procedure at fault? A point of contention amongst developmental researchers is the effectiveness of the looking time paradigm. There are several types of experimental procedures involving preferential looking; the most commonly used involve habituation. Typically the infant is habituated to a stimulus and then successively shown two test displays; one is similar to the habituation display and another is apparently novel. Inferences are made about the perceptual and conceptual level of an infant based on preferential looking.

Haith (1998) points out that it is difficult to know if preferential looking during the test phase is based on familiarity during habituation, or familiarity from previous experience with the world. Reasons for preferential looking can be based on perceptual or conceptual knowledge, or a mix of both – “many factors affect looking, including variations in the perceptual dimensions of objects and people, familiarity, novelty, recency, predictability, and the time lapse between stimulus exposures” (p. 170).

Aslin and Fiser (2005) argue that the methodology of grouping infants and interpreting results based on statistics can be misleading if a confirmatory scientific approach is taken.

Experiments often involve groups of infants because the typical attention span of an individual is short. Group results seem to offer explanatory power, but the actual variable that caused the result remains unknown in a binary yes/no test. If the expected outcome is not predicted in advance, the researcher can create a confirmatory explanation for a positive result, but treat a negative result as a problem with the experimental design (Aslin & Fiser, 2005).

The violation of expectation (VoE) procedure illustrates this point. The VoE procedure replaces the familiarity/novelty preference of the original habituation procedure with a possibility/impossibility preference. Rather than testing how well the infant has learned a stimulus, VoE attempts to test for pre-existing conceptual knowledge. The issue with this, as Aslin and Fiser point out, is that a positive effect can be interpreted as conceptual knowledge while a negative effect can be interpreted as a false negative requiring a tweaking of variables.

### **2.3.2 The Assumption of Object Persistence**

Baillargeon (2008) borrows from Spelke's original definition of an object and develops it further. The object persistence principle requires continuity, cohesion, plus persistence through property change. The ideas inherent to object persistence have some cross-over between the philosophical and the psychological literature. Scholl (2007) identifies the three most foundational philosophical, object-related themes in humans as being spatiotemporal continuity, persistence through property change, and cohesion. Clearly, these align closely with Baillargeon's object persistence principle which is often attributed to young infants.

Fields asks the question "How does an object become subject to the persistence principle?" and consequently rejects object persistence from a neuroscientific perspective. He provides evidence challenging the requirement of innate object knowledge with regards to categorisation, continuity, and cohesion (Fields, 2013). Fields argues that the object categorisation process takes significantly longer than does the process to set up the "object file" – the pointer to which object properties can be bound. Therefore, any principles of persistence dependent on something being labelled as an object would occur

after the object file has been created, making persistence the result of visual perception, not a prerequisite for it (Fields, 2013).

In regards to continuity and cohesion, Fields uses the example of a point-light walker to demonstrate that the timeline of object-file creation doesn't fit with the requirement of object persistence. A point-light walker involves the biological animation of an animate walker. The point-light walker is only perceptible while there is motion. A static frame of the animation appears as a random collection of points. Newborn infants provide differing responses to variances in point-light walker displays (Bardi, Regolin & Simion, 2011), and six-month-old infants detect differences in a walker's direction (Kuhlmeier, Troje & Lee, 2010). Using evidence from adults (which is applicable if the infant also has the same capability), Fields shows that the time it takes to identify a point-light walker is less than the time it would take to categorise the walker as a physical object. Therefore, continuity and cohesion of the points occurs prior to the categorisation of the point-light walker as an object (Fields, 2013).

The point of this argument is not to deny that objects have properties. Rather, Fields is pointing out that assuming innate object knowledge is theoretically vacuous and results in issues like object segregation and re-identification being avoided or ignored.

## **2.4 The Relation to Mental Representation**

With an understanding of the fundamental issue inherent to many theoretical and computational models of development, it is appropriate to next look at why it became an issue in the first place. This discussion requires a historical look at one of the fundamental questions in developmental research: what is the origin of mental representation? Mental representation typically includes any mental imagery that is not directly related to the current input from the senses. This includes memories, ideas, and knowledge.

Piaget was a proponent for the emergence of representation from action with his pioneering work in developmental psychology. Piaget's (1952) constructivist approach sees mental representation as an emergent property, arriving after the association of perception and action. Therefore, the infant is initially a purely reflexive organism. On

the other hand, nativists including Baillargeon, Spelke, and others contend that representation is innate.

### **2.4.1 A Brief History of Developmental Research**

According to Piaget, motor abilities are essential to cognition, and therefore, testing motor skills is a good way to ascertain the cognitive level of an infant (Piaget, 1952). Piaget believed that the progression through infancy and childhood was marked by distinct stages in which certain skills and concepts were learned, with each stage building off the previous one. The first is the sensorimotor stage (from 0 – 2 years old) in which the infant has very little cognitive ability and instead primarily operates by forming relationships between perception and action; Piaget referred to these relationships as schemas. Initially the infant is only capable of reflexive motor actions, but over time reflexes turn into controlled motor actions such as the movement of the hand to the mouth. These basic schemas are used as a foundation for ever more complex interactions with the environment.

Piaget (1952) proposed that at an early age object knowledge must be limited. Object segregation and continuity are problematic and this is demonstrated by failures in reaching tasks. One of Piaget's central claims was that young infants perceive the world literally. An infant can perceive an object only so long as it is in sight. Once the object is out of sight it disappears from existence. Piaget noticed that between four and six months of age infants begin to notice their ability to affect change on the world based on their actions which leads to the notion of cause and effect and related goal directedness. Infants begin to perform actions with the intention of them leading to a certain consequence. At around eight months of age an infant is able to represent an object that is hidden from view. This argument is supported by an experiment in which an infant watches as a toy is hidden underneath a blanket. If the infant does not search the blanket for the toy, it must be because the infant perceives the toy as no longer existing. In contrast, if the infant does search for the toy, then a mental representation of the toy must have been created and the object is perceived as continuing to exist while out of sight.

Over the past few decades Piaget's constructivist theory of developmental progression has been challenged by theories that involve some innate object knowledge (Spelke, 1994;

Baillargeon, 2008). The experimental procedure of Piaget's hidden toy test was fundamentally altered so that object permanence could be tested without the requirement of motor action. A new testing paradigm was devised in which conceptual knowledge is inferred based on preferential looking (Spelke, 1985).

As a result of the new experimental procedure, a vast amount of evidence has emerged that contends that Piaget may have underestimated the abilities of very young infants (Kellman & Spelke, 1983), and that an infant's cognitive and motor abilities are not directly related. This evidence has been used to argue that young infants fail to search Piaget's blanket because they lack the motor abilities to do so, not because they don't understand that the toy continues to exist. The problem with this argument is the implicit assumption that the researcher is still testing the same concept in both situations (Haith, 1998). In the case of object permanence, by changing the experimental procedure the definition of the concept was also necessarily altered. The sensorimotor theory requires that interaction with an object precedes mental representation. This allows the infant to learn about object properties such as solidity and segregation. In contrast, the looking time paradigm requires that a representation of an object can be created prior to any physical contact. This in turn requires innate object knowledge.

Nativists choose to look at this from another perspective. The looking time evidence seems to fit nicely with the hypothesis that infants have conceptual ability from very early on and so a confirmatory scientific approach is taken. Rather than question the reason for the results, the stance is taken that innate object knowledge fits as an explanation and so must be the case. The result is that conceptual level knowledge is often attributed to infants based on looking time alone. Additionally, because looking time experiments can be employed on infants as young as a few days old, there is little opportunity to disprove the findings empirically. However, as will be shown, re-interpretations of the experimental evidence have been effective in posing the question.

In spite of the theoretical issues, nativism succeeded in questioning Piagetian theory to the extent that it is no longer the predominant theory in developmental research. Empiricism has emerged in its place as a common ground between the two. Empiricists share the notion that knowledge is primarily learned, but unlike Piaget's firm stance on

emergent representation, empiricists exist on a spectrum between allowing absolutely no innate knowledge and allowing only enough to bootstrap the learning process. The so-called middle-ground empiricist approaches attempt to explain infant development using limited domain-specific modules (e.g. Spelke & Kinzler, 2007; Newcombe & Learmonth, 1999).

### **2.4.2 Representational Emergence**

An important part of Piagetian theory is overlooked by both nativism and empiricism. The two are supposedly mutually exclusive and at opposite ends of the nature vs. nurture debate; however, as Spencer et al., (2009) and Allen and Bickhard (2013) argue in regards to representation, nativism and empiricism are two sides of the same coin. Both rely on a foundation of representation, as opposed to the emergence of representation proposed by Piagetian theory. The problem is that a representational foundation cannot be explained by non-representational means. A representation requires an external interpreter to provide meaning. The external interpreter then requires its own interpreter. This introduces an infinite regress until a foundation of representation is reached (Allen & Bickhard, 2013).

The issue of representational emergence is indirectly relevant to the current research. It is a foundational issue. Developmental researchers often assume a representational foundation and also include content for that foundation, which in the case of visual perception is object knowledge. Essentially, the assumption of innate object knowledge is symptomatic of the assumption of innate representation. As an example, if an infant is attributed with the ability to mentally represent a situation involving a toy being hidden underneath a blanket, the infant is also implicitly attributed with the ability to separate that toy as a distinct entity from the blanket. Without the latter, the representation provides significantly less useable information to the perceiver.

### **2.4.3 The Conflation of Representation and Object Knowledge**

#### **Meltzoff and Moore's Representational Persistence**

The conflation of object knowledge and innate representation is evidenced in Meltzoff and Moore's (1998; 2001) model of the development of object representation, identity

and permanence. Their model was built in response to the disparity between the Piagetian and nativist developmental timelines outlined previously. They side with the Piagetian view that a perceptual rather than a conceptual explanation is more likely. However, they resolve the issue by claiming that infants possess representational persistence – a mechanism that allows the infant to track spatiotemporal object information. This places their theory somewhere within the empiricist spectrum. Representational persistence allows the infant to follow an object on a temporarily occluded trajectory without necessarily needing to know the featural properties of the object; therefore, the authors can reject the requirement of early object permanence in exchange for object representation.

In their model, object permanence and object representation are understood separately. They argue that object permanence requires identity and representation, but that the reverse is not also true. Therefore, an infant can represent an occluded object without needing to understand that it exists in the real world. Representational persistence will keep track of the location of the object and use this information to later re-identify it as the same individual. It seems that Meltzoff and Moore are essentially replacing permanence of featural information with permanence of spatiotemporal information. However, they see the difference as being that the former requires conceptual knowledge of a thing continuing to exist while out of sight, but the latter requires no such understanding. Further, they identify two types of representation. Previously perceived representation occurs when an object is temporarily occluded while moving along a trajectory. If the trajectory can be predicted, then the object can be re-identified. In contrast, if the object were to stop behind the occluder, then the representation would be lost.

The assumption that infants have an innate understanding of object knowledge to support representational persistence is fundamentally important to this model. Meltzoff and Moore (1998) contend that “infants are evolutionarily prepared for interacting with and representing objects in a steady-state world” (p. 219) by reasoning that “human perceptual systems are adapted to perceive and interact with ‘middle-sized objects’ lying somewhere between atoms and heavenly bodies” (p. 220). Additionally, their model states that the

first level of representation allows neonates to set up representations of objects and events after brief encounters (Meltzoff & Moore, 1998).

### **Mandler's Image-Schema Theory**

Mandler's Image-Schema theory (1988; 1992; 2012) is described as being of an architecture distinct to the typical nativist and empiricist approach of including innate domain-specific modules. Instead, a single domain-general mechanism is responsible for generating concepts from perceptual data. Mandler justified a domain-general mechanism based on her arguments against mental representation emerging from sensorimotor schemas, for which Piaget did not provide a clear transition (Mandler, 1988; 1992). According to Mandler, the procedural memory underlying motor action is not accessible to consciousness, and thus, the emergence of mental representation cannot be explained.

Mandler (2012) contends that Perceptual Meaning Analysis (PMA) is responsible for translating spatiotemporal information into conceptual primitives, which can in turn be grouped into an image-schema. She has described the conceptual primitives as being abstract spatial entities. PMA is responsible for taking visual information and interpreting it as one or more of these primitives.

Mandler (2012) identifies a collection of conceptual primitives that are the building blocks of conceptual knowledge. As an example, a set of conceptual primitives can be used to differentiate animate and inanimate objects.

*The PMA output primitives needed to conceptually differentiate animals from inanimate things are the following: **THING** refers to any perceptually bounded cohesive object. **PATH** refers to any object's **MOTION** trajectory through space. **START PATH** refers to the onset of **MOTION** along a path and **END PATH** to its cessation. **CONTACT** refers to one object touching another. **LINK** refers to a variety of contingent interactions between objects as when a hand picks an object up, back and forth interactions of people, or between paths as when one object chases another. (p. 429)*

The collection of these primitives that make up an image-schema for an animate thing such as an animal are: **THING**, **START PATH**, **CONTACT**, and **LINK**, which translates to a thing that starts moving without contact and interacts with other objects.

Mandler (1992) argues that the spatial structure of image-schemas avoids many of the general issues of empiricism; image-schemas inherently have meaning because they are not images. It seems however, that the issue remains: how exactly does an image of an object come to be a spatial concept with meaning? Müller and Overton (1998) point out that:

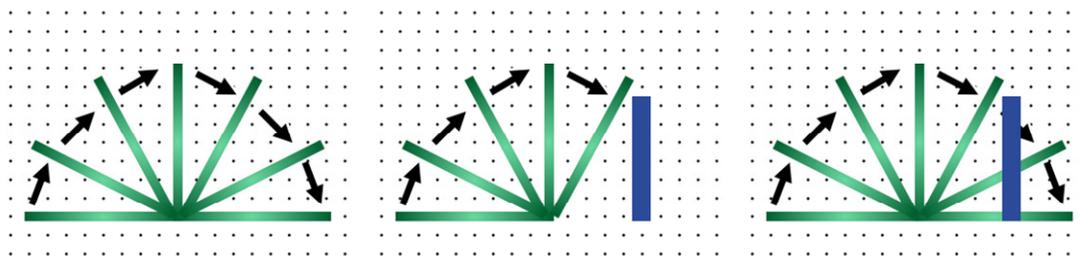
*The idea of a triangle includes the right-angled, acute-angled, and obtuse-angled triangle, but an image cannot be all these at once. Only with the development of the organizing process of knowing, and the comprehension of hierarchical relationships between concepts, does an image of a particular serve as a signifier for general and abstract concepts (p. 80).*

In regards to object knowledge, this argument is applicable to conceptual primitives such as contact. What constitutes two objects coming into contact with each other? How are two objects defined as individuals? As Müller and Overton call attention to, these questions are answerable with categorical knowledge of different objects. But prior to this understanding, any definition is ambiguous and difficult to defend as being basic domain-general knowledge.

The conflation of object knowledge and representation in image-schema theory is clear. Mandler describes a conceptual primitive THING as referring to any perceptually bounded cohesive object. The domain-general mechanism that is responsible for translating sensory input into a THING is not explanatory from a bottom-up perspective.

## 2.5 Methodologically Related Literature

### 2.5.1 Re-Interpretation of the Drawbridge Experiment



*Figure 1: Drawbridge Experiment. Left: Habituation display. Centre: Occluder trajectory ends at 112° (possible event test display). Right: Occluder trajectory continues past 112° (impossible event test display). Adapted from Baillargeon (1987).*

The related literature primarily consists of re-interpretations of nativist experiments that provide a perceptual rather than a conceptual explanation for results. The drawbridge experiment tests the perceived permanence of a temporarily occluded block and seems to be the most re-interpreted experiment in the literature. The drawbridge experiment involves an occluder that moves in a 180° arc just like a drawbridge in front of a castle would do (Baillargeon, Spelke, Wasserman, 1985; Baillargeon, 1987). The infant is habituated to the display pictured on the left of figure 1, however not from the perspective of the displayed image. Rather, the infant is positioned on the left of the display and views the occluder moving away, similarly to how a person would view a drawbridge closing if standing directly in front.

The infant is then presented with two test displays, both of which begin with a block positioned so as to cause a collision during the arc of the occluder. The arc of the occluder in one test display ends at 180° without a collision. An adult perceives this as the occluder squashing the block, an impossible event. The arc of the other test display ends at 112° which an adult perceives as stopping because of a hidden contact with the block. Baillargeon makes the assumption that infants who look longer at the test display with the complete 180° arc must mentally represent the block as existing behind the occluder and become confused when the block doesn't cause a collision. Infants of five months, and some infants of three and four months of age looked longer at the seemingly impossible event. Consequently, Baillargeon contends that these infants maintain a representation of

the block after it is hidden from view. The infant also understands that the block is a solid object that should intersect the path of the occluder, and that the absence of a collision is unexpected.

Rivera, Wakeley and Langer (1999) looked at the experimental procedure and noticed that the impossible event also happened to consist of  $68^\circ$  more motion than the possible event; a clear perceptual advantage over the possible event. They removed the habituation trials from the experiment and found that infants looked preferentially at the  $180^\circ$  test display. They then repeated the original experiment but removed the block from the  $180^\circ$  test display so that this was no longer an impossible event. Infants again looked preferentially at the  $180^\circ$  test display. Rivera et al. (1999) concluded that the increased motion was the cause of the preferential looking.

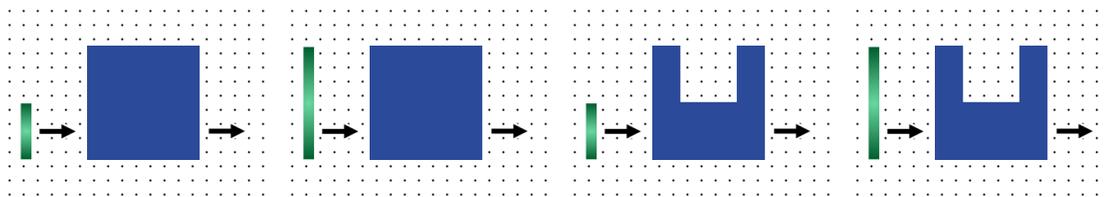
Bogartz, Shinsky and Schilling (2000) repeated the original experiment using an event set  $\times$  event set design to test the effects of the habituation paradigm of the findings. This design requires splitting the infants into groups and habituating each to either the original habituation display, the possible test display, or the impossible test display. Each group is then tested using the two displays to which they were not habituated. This results in data for every possible combination of habituation and test. Bogartz et al. (2000) showed that the looking time results of the experiment were caused by familiarity/novelty preferences, rather than by a possible/impossible event distinction as was originally interpreted. Cashon and Cohen (2000) used a similar experimental procedure but added an event consisting of the occluder stopping at  $120^\circ$  without a block. They habituated groups of infants to one display and tested on all four. Like Bogartz et al. (2000), they found that perceptual level familiarity/novelty was the cause of preferential looking.

All three analyses show that variations in either familiarity with a stimulus, or visible perceptual properties are sufficient to reverse the original findings, and therefore the interpretation. The breadth of criticism directed at the drawbridge experiment is perhaps a reflection of a flawed methodology. However, the criticism also supports the more fundamental argument against nativist assumptions of infant cognition. Careful consideration of the experimental procedure must be given before attributing infants with conceptual knowledge.

Baillargeon responds to these criticisms by finding issue with the methods used (Baillargeon, 2000). While it is true that the methods were different – typically because the objective was to find fault with the original method – this only provides a limited defence. If the infant is actually demonstrating conceptual ability, the result should remain similar given minor variations in the stimulus. Baillargeon also defends the experiment on the grounds that dozens of others have since confirmed her results using the same or similar methods. In reply to Bogartz et al., (2000), Baillargeon argues that an issue with one experiment should not discredit the work as a whole. This is a fair request, however even without the breadth of similar criticism, it at least demands further study.

In addition to a general critique of nativism, these re-interpretations are included in the literature review to showcase the type of re-interpretation currently available. Unlike these, the current research is not concerned with questioning individual experiments by changing variables in the experimental procedure. Instead, this work is broader and more theoretical. The interpretive process is similar in that conceptual knowledge is not assumed, and instead a perceptual bottom-up approach to interpretation is taken. However, the objective is to theorise the effects of the bottom-up approach on different experiments rather than to empirically confirm them on an individual.

## 2.5.2 Re-Interpretation of the Short and Tall Object Experiment



*Figure 2: Short and Tall Object Experiment. Left: Short object habituation display. Centre left: Tall object habituation display. Centre right: Short object test display. Right: Tall object test display. Adapted from Baillargeon & DeVos (1991).*

The short and tall object experiment again tests object permanence in infants. It involves a short and a tall object, and an occluder. The infant is habituated to both objects moving on a trajectory from one side of the display to the other behind the occluder, as depicted on the left of figure 2. The test phase involves the same stimulus except that the occluder is replaced with another that has a window cut out of the centre, as depicted on the right

of figure 2. The trajectory of the short object is equivalent to the habituation display. In contrast, the tall object reaches the edge of the occluder, but doesn't appear in the window as it passes behind. It then reappears on the opposite edge of the occluder. To an adult perceiver, the trajectory of the tall object is impossible. It should have appeared in the window during its trajectory behind the occluder. Baillargeon found that three and a half month old infants looked preferentially at the impossible event, and interpreted this as being because they represented the height and trajectory of the object and were surprised when it didn't appear in the window of the occluder (Baillargeon & DeVos, 1991).

Meltzoff and Moore (1998) use their theory of representational persistence to re-interpret this experiment. They point out that if the infant can extrapolate the trajectory of the object, then the surprise should be apparent at the instant that the object doesn't appear in the window. The infant shouldn't need to wait for the reappearance of the object to be surprised. As previously discussed, Meltzoff and Moore do not require that the infant is able to represent the object as continuing to exist in the real-world. Instead, they require that the spatiotemporal position of the object is tracked.

The infant no longer requires a conceptual understanding of object permanence. However, the assumption remains that the infant is able to maintain a representation of the objects while out of sight. This requires object knowledge about continuity and occlusion. Essentially, the definition of representation and object permanence has been changed again. It will be recalled that Piaget's original definition included both representation and object permanence. Meltzoff and Moore argue that only object identity is necessary, and that the understanding that the same object continues to exist in the world is unimportant. Instead of a high level mechanism for maintaining an object's featural information all that is required is a low level mechanism to track object location. The justification for this mechanism is that humans are innately endowed with an understanding of the "steady-state" world – a world full of average-sized objects (Meltzoff & Moore, 1998).

Meltzoff and Moore's model of representational persistence is an example of the empiricist approach to problem resolution. A more plausible interpretation of the empirical data, results in the problem being moved to a lower level of abstraction and

ignored. The fundamental issue – the origin of object knowledge – inherent to the nativist interpretation is still applicable, albeit less overtly.

## **2.6 Computational and Algorithmic Analysis**

The methodologically related literature includes re-interpretations of experiments using variations of the original method, or conceptual theory. Neither of these approaches is a viable method for investigating the development of early object knowledge because both require specific object primitives.

Marr (1982) provides the framework for computational and algorithmic analysis that can be used to interpret a scene without any preconceived assumptions. Marr contends that vision was an information-processing task that should be analysed at the computational, algorithmic, and implementation level. Analysis at each level serves a different purpose.

- The computational level: The problem is described generally and constraints on a solution to the problem are specified.
- The algorithmic level: The input and output information associated with the solution are specified, and an algorithm for transforming the input to the output information is designed.
- The implementation level: The neural structures responsible for solving the problem are identified.

Although the three levels are designed to build from computational to implementation, they can be considered independently (Marr, 1982). Of the three, the computational and algorithmic levels of analysis are relevant to the current research objective. The implementation level is used to find a neuroscientific explanation for the algorithm and so is outside the scope of this research.

Marr (1982) uses the algorithmic level to describe the transformation of retinal images into three-dimensional representations of objects. This includes the mechanisms and representations necessary to go from luminance intensities, to an orientation and depth map, and finally to a hierarchical structure of three-dimensional models.

Gibson (1986) argues that rather than requiring representation, visual information could be directly picked up and used by the perceiver so long as a reason exists to do so. The “ecological” approach assumes that vision is a tool that enables animals to accomplish goals such as finding food and avoiding predators – a claim that seems evolutionarily plausible. However, Gibson’s proposals for the algorithmic level descriptions of these direct perceptual mechanisms were often unclear (Glennester, 2002).

There are similarities between the debate of Marr and Gibson, and the empiricist and constructivist debate. Marr believes that vision is an information-processing task and so can be reduced and reverse engineered to a representational foundation. Gibson believes, as Piaget did with young infants, that perception is based on goal-directed motor action. The development of an interpretive approach will borrow from both of these accounts of visual perception.

## 2.7 Summary

The literature review has attempted to highlight three points that justify the knowledge gap for this research.

- The assumption of innate object knowledge is symptomatic of the assumption of innate representation, and therefore prevalent throughout the developmental literature.
- Methodologically related literature is limited, making this research exploratory. However, this research follows logically from existing issues.
- A non-conceptual approach to interpretation is required to avoid making assumptions about innate object knowledge.

Computational models of developmental theory seem to be preconfigured with innate assumptions about objects. These assumptions are knowingly or unknowingly transferred from the developmental research that they are built to support.

The assumption of innate object knowledge is symptomatic of the assumption of innate representation, and therefore prevalent throughout the developmental literature; however, despite its prevalence, debate surrounding the effects of this issue is lacking. Perhaps this is because of the difficulty involved in framing a developmental theory or experimental interpretation without using some reference to the term “object”; by assuming some object knowledge is innate fundamental questions can be ignored including: what does it mean to re-identify an object after an occlusion event? And how is an object segregated from its surroundings as an individual? It may be that these questions are ignored because the answers seem obvious: object segregation is the demarcation of objects based on colour and texture. The purpose of this research is to think more deeply about how researchers come to these conclusions. In the real world an object is never only a single colour and texture, in which case it follows that assuming no other object knowledge this information should have little impact on object segregation.

The methodologically related literature is concerned with questioning specific experimental procedures, or the looking-time paradigm itself, rather than with the theoretical issues inherent to innate object knowledge. Re-interpretations of experiments

show empirical data cannot always be trusted, and that experimental procedures must be carefully constructed. The underlying point of the review, however, is to show the difference in perceptual and conceptual interpretations of the same data. Rivera et al., (1999), Bogartz et al., (2000), and Cashon and Cohen (2000) show that a non-conceptual interpretive approach is likely useful in order to avoid making assumptions about innate object knowledge.

However, the re-interpretations reviewed are not designed to theoretically explain the development of the object knowledge involved. It seems that a novel interpretive approach is required. Marr's computational and algorithmic levels will be used as the framework for a bottom-up interpretive approach. The computational level aligns closely with the identification of assumed object knowledge, while the algorithmic level is used as the basis for the re-interpretation of the events of an experiment. Essentially, the bottom-up interpretive approach is a method for the computational analysis of the information important to perceptual processing.

The bottom-up approach, like Piagetian theory, begins with the learning of reflex motion in response to a stimulus and builds up from there. Therefore, it doesn't assume any of the properties of object persistence. Rather than automatically treating objects as individual entities, the bottom-up interpretation of a scene is a literal view of visual perception, similar to what you would expect of a computer. In place of an object, it might be sensible to begin with Gibson's (1986) notion of a surface to describe regions of a scene. In contrast to an object, a surface is a literal view containing only the information available from the light arriving at the retina. Gibson (1986) explains:

*According to classical physics, the universe consists of bodies in space. We are tempted to assume, therefore, that we live in a physical world consisting of bodies in space and that what we perceive consists of objects in space. But this is very dubious. The terrestrial environment is better described in terms of a medium, substances, and the surfaces that separate them. (p. 16)*

## 3 RESEARCH DESIGN AND IMPLEMENTATION

### 3.1 Introduction

Typically when designing a research project, selecting a method and justifying that selection is fairly straightforward. There are a limited number of tried and tested options available to help answer a research question. This thesis is unique in that the data analysis required does not align with any of the existing options that I have come across. Instead, the design has evolved from the combination of different approaches; it culminates in something unique to the objectives of this thesis, but which could be modified for use elsewhere.

Given the ambiguity associated with this research design, it is important to make design decisions in a consistent and logical way. Michael Crotty (1998) provides a process that supports the selection and justification of the four elements of a research design. This process begins with the selection of a method, and progresses with justification based on methodology, theoretical perspective, and epistemology. Crotty (1998) defines these elements as follows:

*Methods: the techniques or procedures used to gather and analyse data related to some research question or hypothesis.*

*Methodology: the strategy, plan of action, process, or design lying behind the choice and use of particular methods and linking the choice and use of methods to the desired outcomes.*

*Theoretical perspective: the philosophical stance informing the methodology and thus providing a context for the process and grounding its logic and criteria.*

*Epistemology: the theory of knowledge embedded in the theoretical perspective and thereby in the methodology. (p. 3)*

According to this process, the data collection and analysis method is selected first. Typically, this might be a survey or interviews, followed by quantitative or qualitative analysis. The data required for this research is unique to the objective and the method is more difficult to define. The objective is to explore visual perception by re-describing infant behavioural experiments using lower level visual primitives than were used

originally. The data is not categorically quantitative or qualitative, rather it could be described as a hybrid. The original interpretations of experiments are generally objective accounts of visual phenomena; therefore, the empirical evidence for most experiments is quantitative. However, these original interpretations are then re-described in a manner that is subjective to the researcher. The researcher cannot know how a stimulus is perceived by an infant and must make educated inferences as to the mechanisms behind perception. Interpreting the mechanisms responsible for the subjective experience of visual perception requires a combination of empirical evidence and imagination.

This chapter begins with an overview of the high level research approach. This is not included in Crotty's scaffolding because justification is typically not necessary with a confirmatory approach to research.

### **3.2 Exploratory Research Approach**

This thesis requires an exploratory rather than a confirmatory approach to research. The confirmatory approach refers to traditional research that attempts to test whether something is consistent with prior assumptions, theories, or hypotheses. The confirmatory approach typically requires a clear research question and hypothesis, and follows a structured process to progress from question to conclusion. The focus can be put on testing one or a small number of variables, and this testing can later be verified by others by following the same process. In contrast, the exploratory approach is focused on investigating a research topic from a unique, or non-obvious perspective (Reiter, 2013). The aim is to develop new ideas and hypotheses rather than to confirm existing ones empirically.

The exploratory researcher typically has more freedom for change as a project progresses. For this reason it is important that strict constraints be put in place so that the project is managed effectively and continues in the intended direction. There is a significant risk that if not managed effectively, this approach to research can result in a broad and ambiguous project that fails to find anything of value. To ensure the value and reliability of findings, the researcher should be as transparent and honest as possible (Reiter, 2013). This becomes even more important when considering that the research output is typically qualitative rather than quantitative; the unprincipled researcher has ample opportunity to

knowingly or unknowingly bias the results. The difficult part of developing an exploratory research method is balancing the constraints that limit scope creep with those that provide the opportunity to discover new knowledge. One way to achieve this is with the ongoing iterative improvement of the research design. Indeed, the research methods developed for this thesis are the result of several iterations of development. It is important to note that the perfect research design doesn't guarantee the discovery of new knowledge. Given the nature of this approach, even if managed effectively this type of research can conclude with limited or no interesting findings. Such is the risk of exploration.

Although the confirmatory approach is overwhelmingly more common in the sciences, it also has limitations based on its philosophy. The reason exploratory research can be a valid option is because research is rarely ever strictly confirmatory (Reiter, 2013). This is seldom discussed because confirmatory research must be quantifiable in order to be of value; however, research generally exists on a spectrum. Karl Popper (2005) tells us that theories cannot be unequivocally confirmed, and instead can only be verified as the best explanation until falsified or until a better one is discovered. This is because any scientific confirmation is the result of thoughts and ideas originating in the mind. It is impossible to know if anything is objectively equivalent to reality. The confirmatory researcher must rely on verified but ultimately unproven theories when developing a research project. The decisions made about which theories to test and use as the basis for hypotheses are often subjective and can potentially introduce bias to any findings as there is typically no self-reflection involved in the design process to justify these decisions. In contrast, the exploratory researcher understands that scientific discovery is constructed by the mind and may not be objectively real because the language we use for explanation is also a construction. As a result, the exploratory researcher makes more effort to understand the foundational biases involved in research and uses this to look at the problem from a unique perspective.

The difference between perception and reality is of significant relevance to this thesis because the general theme is the interpretation of mechanisms involved in visual perception; a topic which is toward the top of the list of phenomena which are difficult to objectively study and prove. The interpretations are also specifically about the language

used by psychologists to explain foundational perception. Concepts such as occlusion and permanence can mean different things to different researchers making universal objectivity difficult to achieve. The exploratory approach is appropriate for this thesis because it allows for a detailed examination of what concepts like permanence mean by breaking them down and thus removing bias.

### **3.3 Research Methods**

As Crotty (1998) suggests, the techniques we plan to use should be described first. This section offers an overview of the research methods. A more thorough description follows in section 3.5.

Perhaps the most unique aspect of this research is that the data to be analysed necessarily relies on the interpretation of observations provided by other researchers. Data collection does not involve primary data in the form of surveys or interviews as is typically the case in qualitative research. Instead, data is collected entirely through a review of the developmental literature, specifically experiments involving visual perception in infancy. Two facets of each experiment are analysed. First, the results of the experiment are used as a foundation, on top of which a new interpretation of a perceptual concept is developed. This results in the new interpretation being grounded, to whatever extent this is possible, in the original empirical evidence from behavioural experiments. Second, the validity of this new interpretation is justified by critically analysing and contrasting it with the interpretation of the original researchers. If the new interpretation is able to account for the empirical evidence without requiring assumed object knowledge then it can be considered a successful finding.

The re-interpretation of experiments individually will result in a limited amount of insight into visual perception as a whole. Each experiment represents a single or small number of perceptual concepts because as is typically the case with the scientific method, an experiment tests a single or small number of variables. More insight is undoubtedly available if these pieces can be assembled into a developmental pathway that shows how the co-operation of lower level primitives can result in higher level concepts. Determining this developmental pathway is the ultimate aim of this research. Unfortunately it is difficult to know ahead of time which experiments to select for interpretation; until

interpretation the manner in which two experiments might relate can only be speculated. Problems inherent to the research method will not be clear until after several experiments have been interpreted and so updates are made iteratively as the research progresses.

### **3.4 Methodology, Theoretical Perspective, and Epistemology**

The research methods must be grounded in an appropriate philosophy of science and the philosophic perspective should be justified through a comparison with other options. Following Crotty's process, the methodology will first be selected and described. There are several methodologies that partially meet the requirements of this research but none that I have found that match the requirements exactly. This is likely because it is difficult to build a generic framework around exploratory research, but also because the research method involves two distinct parts.

Firstly, the literature must be reviewed in order to find viable developmental experiments on object knowledge in infancy. This requires its own data analysis phase in which methodologies, results, and interpretations are reduced into collections of similar object concepts and their possible perceptual connections. These connections are vital to meet the objective of this project because without them it is much more likely that the end result will be a collection of isolated re-interpretations of experiments with no clear developmental progression. The second phase involves the re-interpretation process itself which will involve using the data gathered in phase 1 to build a model of the development of early object knowledge.

The separation of the research method into two phases is strictly to bring lucidity to an otherwise protracted design. It may be more accurate to describe the first phase as prerequisite data analysis for the principally important second phase. The second phase being the only one that results in any kind of directly relevant findings. However, the first phase remains important for removing potential bias from the selection of secondary data, and is therefore still described in depth.

#### **3.4.1 Methodology for Phase 1 – Data Search and Initial Analysis**

Often with exploratory qualitative research, a grounded theory approach can be taken. In classical grounded theory there should be no preconceived expectations as to the result of

the research, all theory should emerge from the data (Glaser & Strauss, 2009). Generally the raw data for grounded theory is coded data from interviews, but observations are also acceptable and observations are essentially the output of experiments of infant development. The issue with grounded theory is that it is a highly structured process because it strives to gain meaningful results from data without a prior research question. In this case, there is a research objective and question, and although the exploratory nature of the research defines them as being merely guidelines, they remain a clear direction to follow. Therefore, grounded theory is an unnecessarily complicated method and methodology for the proposed research.

General inductive analysis is a similar but less structured version of grounded theory (Thomas, 2006). It allows the researcher to focus on the most important/interesting concepts and themes that emerge from the data rather than systematically reviewing all of them. General inductive analysis is also more appropriate than grounded theory because of the underlying theoretical nature of developmental research. Many ideas will be based on interpretations of observations of infants which are difficult to validate in comparison to interpretations of the reactions of children or adults. This means that interpretations from the existing literature may unavoidably rely on developmental theory and may not be exclusively derived from raw experimental data. The primary purpose of the general inductive approach is to allow research findings to emerge from the frequent, dominant, or significant themes inherent in raw data, without the restraints imposed by structured methodologies (Thomas, 2006).

The general inductive approach closely fits the data analysis required for the first phase of the current research because it allows the raw data, in this case the experiments of infant development, to be collated to find themes related to object concepts as desired by the researcher. In contrast, the grounded theory approach would be more suitable for finding themes in general.

### **3.4.2 Methodology for Phase 2 – Re-Interpretations of Experiments**

The second phase of research is more difficult to ground within a specific methodology. It involves using the data collected in phase 1 to identify the input and output variables for different object concepts. This requires an interpretation of the conceivable

relationship between the input and output variables, and results in theoretical object knowledge. The re-interpretation of experiments is subjective to the researcher. It is the researcher's interpretation of what it might be like if visual perception were devoid of meaning. This is not a typical type of research and part of the reason why it is described as exploratory.

Although this is similar to a subjective account of an actor's experience, it is actually a simulated subjective account, thus it is not phenomenological because there is no actual experience. There are some basic similarities with other qualitative methodologies but none (that I have found) that match completely and so it seems more logical to define it as an exploratory qualitative approach to research.

Reiter (2013) suggests that a dialectical approach is appropriate for exploratory research. Dialectics involves contrasting opposing ideas to find a middle ground. It includes three components, thesis, antithesis and synthesis; where the thesis is the idea, the antithesis is the opposition to the idea, and the synthesis is the eventual compromise. Dialectical thinking allows for new insights to be discovered through the comparison of qualitative data and ideas. This methodology fits nicely with the perceptual, conceptual interpretive dichotomy on which this research is based. The original interpretation of the experiment is compared directly to the re-interpretation.

### **3.4.3 Theoretical Perspective and Epistemology**

The purpose of the theoretical perspective – “the philosophical stance that lies behind our chosen methodology” (Crotty, 1998, p. 7) – is to acknowledge the assumptions inherent to the chosen methodology or methods, and to attempt to ground them. The raw data – the empirical evidence from infant experiments – is quantitative statistical data; however, as shown in chapter 2 the results can often be contradictory given slight variations in experimental procedure. The assumptions relate to the interpretive bias inherent to developmental research.

There are several possibilities for a theoretical perspective but the common two are positivism and interpretivism. Positivism is about proving or disproving an objective reality through hypothesis and experimentation. This approach typically aligns with

quantitative research methods. In contrast, interpretivism acknowledges a reality that is subjectively constructed and is about investigating a phenomenon through qualitative data collection and analysis. Positivism cannot ground the aforementioned assumptions because the operation of mental representation cannot be experimentally tested. The theoretical perspective may have been positivist if the research objective was to make statistical inferences about the collective results of each experiment. Instead, however, the objective is to re-interpret the experiment from the perspective of the naïve infant. Therefore, the researcher acknowledges that there must be a constructed reality, and in so doing, accepts interpretivism as the theoretical perspective.

Finally, the epistemology inherent to the theoretical perspective can be described. The objectivist epistemology holds that truth and meaning reside in the object, while the constructionist epistemology holds that truth and meaning reside in the mind. As is typically the case with an exploratory approach, this research is constructionist because it is the human interpretation of phenomena. This research is ultimately about describing the development of meaning as it exists in a mental representation. The stance taken is that there can be no meaning without the mind, and therefore, it aligns closely with a constructionist epistemology.

## **3.5 Research Implementation**

### **3.5.1 Data Search Protocol**

Data collection begins with the selection of an appropriate experiment for study. A search protocol is required to ensure that data collected remains in scope. The search protocol is borrowed from Kitchenham (2004) who provides a framework for a systematic review of computer science literature and highlights the importance of a protocol to maintain consistency. The systematic review is generally used to gain insight from the collation of a range of empirical evidence which is collected with a comprehensive literature review. This will not be appropriate for the current research; however, the search protocol is useful in order to find experiments related to the research objective.

The data analysis will be made on the type of experiment rather than on every single variation of the same type. Studies using the same or very similar experiments can be

removed at the researcher's discretion. Both the breadth and depth of this review is important. The literature search must be comprehensive in order to determine all current object-related experiments of development. Some experiments included in articles related to developmental psychology may not be described in enough detail for an accurate evaluation to be made later in the analysis. Unfortunately, the judgement as to whether the article describes the experiment in enough detail is difficult to make at this stage of the process; however, the chance of successful data collection can be improved by including additional articles that detail the underlying mechanisms of the experiment.

<b>Inclusion Criteria</b>	<b>Exclusion Criteria</b>
<p>A study is primarily about the early development of object recognition and/or object reasoning in infants in any context.</p> <p>Rationale: The research question is directed towards investigating object permanence; however, evidence from other concepts may be important to development and should also be included.</p>	<p>A study is primarily about any other type of infant development.</p> <p>Rationale: While evidence of other types of development may involve similar low level mechanisms, they will also likely consist of significant differences that will make arguments difficult to form.</p>
<p>A study is published in an academic journal or book chapter.</p> <p>Rationale: All evidence should have been published with some level of peer review to ensure quality.</p>	<p>A study is not published in an academic journal or book chapter.</p> <p>Rationale: Evidence from sources including magazine articles and video presentations may not be of a high enough quality to provide value to an argument.</p>

*Table 1: Data Search Protocol*

### **3.5.2 Data Analysis**

The desired output of the proposed research will be a model of perception that offers new insight to some of the object knowledge related issues outlined in chapter 2, and others that may be found during the course of the research. The output will result from several iterations of the data analysis phase. A set of perceptual primitives will be defined and/or refined for each iteration. These primitives will provide the criteria to assess each experiment and will be updated depending on the results of the prior iteration, ultimately

improving the final result. As the development of these primitives will be iterative, the final version of these could also be included as an output which would provide more transparency to the final results.

The data analysis methods are borrowed from the qualitative data analysis methods of Miles and Huberman. These include methods for data reduction, data display, data comparison, and conclusion drawing (Miles & Huberman, 1994). These methods allow the researcher to form a more objective argument on research questions that rely on the interpretation of many data sources. Elements of the general inductive method of coding are also used to understand the meaning of interpretations in a broader context. This is achieved through a close reading of the text and the revision of code categories (Thomas, 2006).

#### ***3.5.2.1 Data Reduction***

The data reduction phase involves gathering relevant data from each primary source and recording it in a database. Each experiment selected during the data collection phase is a primary source. Firstly, the primary source will be reviewed to determine quality. Next, data will be extracted and coded based on the research topic, methodology, empirical evidence, interpretations, limitations, and results of the review. The classification system will be adjusted and refined depending on requirements. The outcome of this phase is that the data recorded in the database will show relationships between similar experiments in a succinct and organised manner for later review.

#### ***3.5.2.2 Data Display and Comparison***

The data display and comparison phase involves inferring patterns from the extracted data by grouping and interpreting related data points. Patterns emerge through the critical analysis of data from primary sources. This critical analysis will include finding similar or contrasting data, and finding changes over time by displaying results from experiments by age. Data display can be achieved by manually transforming the data into network diagrams that map relationships between nodes, and charts that present the amount of information supporting each data point.

### ***3.5.2.3 Conclusion Drawing***

The patterns found in the preceding stage are used to draw conclusions and arguments. Conclusions in this context will consist of a ranking of experiments that show potential in regards to further data analysis. Conflicting patterns will be resolved by looking at the amount and the quality of evidence for each. Some ideas in developmental research are theoretical so in the event that a reasonable decision cannot be made between two equally relevant but conflicting patterns, the conclusion will be based on both.

### ***3.5.2.4 Experiment Re-Interpretation Protocol***

The selection of an experiment for re-interpretation is based on the prior data analysis. The selection criteria remains relatively unconstrained. The relationships between different experiments support the relevance of an individual selection by improving the likelihood that later re-interpretations will relate in some way. This improves the probability that the research will not culminate in several isolated re-interpretations that offer no distributed insight.

There are several criteria that are mandatory for the selection of experiments for re-interpretation:

- The topic of the study should be relevant to the research objective. This includes object knowledge related to motion, occlusion, permanence, etc.
- The experimental procedure should be interesting enough that it demonstrates different results for several age groups.
- The researcher's interpretation of the experimental results assumes object knowledge without explicit justification.

The re-interpretation of an experiment begins with a dissection into a history of events. During the first iteration of data analysis, the experiment was broken down into conceptual primitives which were similar to those described by Mandler in her image-schema theory (Mandler, 2012), and discussed in chapter 2. At the time, this seemed appropriate because the spatial analogue nature of the conceptual primitives avoided the issues of initial object knowledge. However, the result of the first iteration of the data analysis phase was that the re-interpretations were no more explanatory than the original

interpretations. Rather than an abstract spatial description of events, a more detailed description of events over time was required.

The second iteration of the data analysis phase changed to a naïve physics-based approach. Using naïve physics, the experiment can be broken down into episodes of time (Hayes, 1985). The episode is subjective to the perceiver rather than being an objective event. For example, an object's movement in the environment may result in an episode for one perceiver, while for another it may be part of a larger event. The important point is that the episode is marked by a start and end event. The advantage of moving from conceptual primitives to this approach is that events occurring within the visual array can be interpreted from a small to a large time scale, as opposed to only as a single unit of measure.

The episodes of time are compared to a list of object knowledge that was the outcome of the re-interpretation of prior experiments. This object knowledge emerges as the bottom-up perceptual inputs required for an explanation of the results of an experiment. When an experiment is re-interpreted there are two possible scenarios. Either the currently known perceptual inputs can explain the results of the experiment, in which case research can progress to the selection of another experiment, or the experimental results cannot be explained without more information. When the latter scenario occurs, the researcher attempts to find new information that could explain the experiment, and failing that the experiment is marked as uninterpretable using current knowledge.

In order to add new object knowledge to the model it must explain an aspect of visual perception, and it should be built exclusively from the foundation of current perceptual primitives. If it is determined that a concept cannot be explained with current knowledge, then assumptions can be added if justified appropriately. The new object knowledge should also be grounded by the experimental evidence from the original study. This means that the new interpretation of perception should match the preferential looking data to the extent that this is possible. Because this data is often contradictory, as discussed in chapter 2, this is again not a mandatory requirement as long as there is an appropriate justification for the new interpretation.

### **3.5.3 Iterative Updates to the Methodology**

During the course of the review and analysis phases further research topics may become apparent. Provided that these contribute a different perspective to the overall discussion, they will be added to the list of research topics and will be included for review in the following iteration.

Typically each iteration will require that the model is reviewed to ensure it is supported by the new information. This includes repeating the re-interpretations of experiments using the new information, or dismissing the experiments entirely.

This project is currently in its third iteration. As discussed previously, the first iteration involved a different interpretive approach using conceptual primitives. To resolve the inherent limitations, the second iteration replaced conceptual primitives with histories and episodes as the unit of analysis for events over time. This iteration required that the model at the time be thrown out because the re-interpretations were determined to lack explanatory power. The third iteration added the constraint that experiments should use an experimental procedure with habituation or familiarisation and successive displays. The rationale was that without habituation, it is too difficult to interpret what the infant might be perceiving. This only required the exclusion of experiments based on the preferential looking procedures that measure pre-existing knowledge.

### **3.5.4 Limitations**

The limitations of this research method are at the same time the advantages of an exploratory over a confirmatory approach to science. The confirmatory approach is powerful when the objective is to empirically ground a hypothesis. However, when an area of study has no foundational confirmatory research, the theory behind that hypothesis is likely to have originated from interpretive research.

The exploratory approach requires that there are few constraints on the re-interpretation protocol. This is necessary and controlled through the data analysis of phase 1, and the iterative development process. Yet, it remains a potential issue and could result in the research progressing in unforeseen ways. Again, this issue could be considered a

limitation, but is also potentially a strength of the approach and allows for new insight to emerge.

The necessary removal of structure from the research process results in another limitation: the increased possibility of interpretive bias. The researcher must remain as objective as possible through a process of constant reflection. This is supported through comparisons between the re-interpretation and the original interpretation of the developmental researcher which serves to justify the bottom-up approach. Additionally, all interpretations will be recorded and later re-evaluated where appropriate to support the development of unbiased conclusions (Miles & Huberman, 1994).

The results of the data search as outlined in section 3.5.1 are not included in the thesis because of the exploratory and iterative nature of the research. The data that was found and ultimately contributed to the findings was collected iteratively and so is difficult to present coherently. Additionally, as discussed in section 3.2, the exploratory research approach can provide useful results without necessarily requiring the meticulous data recording required by the confirmatory approach. The data search protocol is used only to find useful data for re-interpretation, rather than as evidence to support an experimental result. With this rationale, the exclusion of the results of the data search should not negatively impact the results of the thesis. Again, this limitation can be attributed to the exploratory research approach and not to a flaw in the design of the research methods specific to this project.

## 4 FINDINGS AND DISCUSSION

### 4.1 Introduction

This chapter consists of the re-interpretation of several types of behavioural experiments, all of which have been influential to the field of developmental psychology. The purpose of each interpretation is to gain further insight into a concept related to object knowledge by describing the experiment using lower level primitives than the original researchers. The interpretations are hereafter referred to as naïve interpretations, both because the bottom-up analysis requires a naïve perspective of visual perception, and also to aid in distinguishing them from the originals. A naïve interpretation requires that we start with next to no understanding at all and try to build concepts up using their component parts. To this end, even basic concepts like velocity, distance, and direction are not assumed as innate knowledge. As discussed in the research design, interpretations involve breaking down observed perceptual evidence into its important component parts and attempting to explain these individually.

To build this model of early visual perception several assumptions are made. Firstly, it is assumed that the newborn infant has an intrinsic preference for a region of the view that changes over time. This doesn't imply any understanding of motion, instead the purpose of this capability is only to divide regions of the view based on salience. There is empirical evidence that demonstrates a preference for moving over stationary stimuli (Volkman & Dobson, 1976; Slater, Morison, Town & Rose, 1985).

It is also assumed that the infant can learn by associating motor actions with sensory information. This requires a system of positive and negative reinforcement based on responses to combinations of perception and action. An unrelated example of such a system involves consuming food for energy. It is unlikely that an infant makes the conscious decision to eat to increase energy levels and avoid fatigue. Instead, the negative feeling of hunger is quelled by the action of eating, giving this action a positive response to that specific feeling. In these interpretations, an innate system of positive and negative reinforcement is integral to the focus of attention on moving visual stimuli within the environment; hence, we can assume that it is innate.

Finally, it is assumed that perception results in a mental representation of the scene. The mental representation initially only consists of a memory of previous events, as opposed to the typical definition which includes any mental imagery including ideas and knowledge – the abstract rules that can be applied to mental imagery. This is an important distinction; the naïve model allows memories to persist over time, but rejects mental representation that involves changes to a memory until these are explained by perceptual mechanisms. Once a mental representation exists, any inconsistencies caused by sensory information that contends with learned object knowledge results in an externally manifested response by the infant (increased looking in the following examples).

The remainder of the chapter consists of the naïve interpretation of four experiments. The experiments have been ordered to emphasise the developmental progression of the naïve perceiver. The most fundamental perceptual mechanisms are described by the first few interpretations and others are developed from that foundation. The research question is directed towards the role of motion in object permanence and many unrelated visual concepts are consequently ignored as unnecessary in this model. Given that the experiments reviewed comprise a range of different ages in early infancy, any assumptions, in addition to those outlined above, will be made and explained on a case-by-case basis.

The discussion of each experiment follows a deliberate structure that is designed to contrast and compare the original findings with a naïve interpretation. Each section begins with a brief summary and visual illustration of the original experimental procedure. The original findings and interpretation of required object knowledge follows and provides a foundation for later comparison. Next, a detailed re-interpretation of the experiment using the emerging naïve model is presented. Finally, a short discussion justifies the naïve over the original interpretation through the resolution of issues inherent to the original.

The interpretation of experiment 1 provides an introduction to the naïve lens with which experiments will be interpreted. Evidence from eye tracking of a spatiotemporal completion task is used to argue for the relative importance of spatiotemporal over featural information. This leads to the definition of the first perceptual mechanism in the naïve

model. Sensorimotor associations are primarily spatiotemporal and rely on a procedural memory of motor actions.

The interpretation of experiment 2 demonstrates the difficulty of the cohesion of surfaces within a scene, a concept that is often overlooked in regards to explanation and assumed innate. Eye-tracking evidence is used again, but this time from a spatial completion task, and in conjunction with the original experiment. The task involves mentally joining two disparate regions of an object. Based on the evidence, an argument is made that infants of different ages are receiving dissimilar sensory input and that observing the important regions of a scene must be learned before the regions can be joined. To achieve this, another perceptual mechanism is defined as the learned cohesion of important but disparate regions of the scene.

The interpretation of experiment 3 describes a perceptual mechanism for the segregation of surfaces based on occlusion. Object segregation is another ability that is often assumed to be innate. However, the reasoning for this is typically based on three-dimensional object manipulation, a factor that is not relevant to the experimental procedures selected thus far that use a computer display to render the stimulus. A spatiotemporal completion task similar to that of experiment 1 is used to demonstrate how surface segregation could operate and be learned in two-dimensional space.

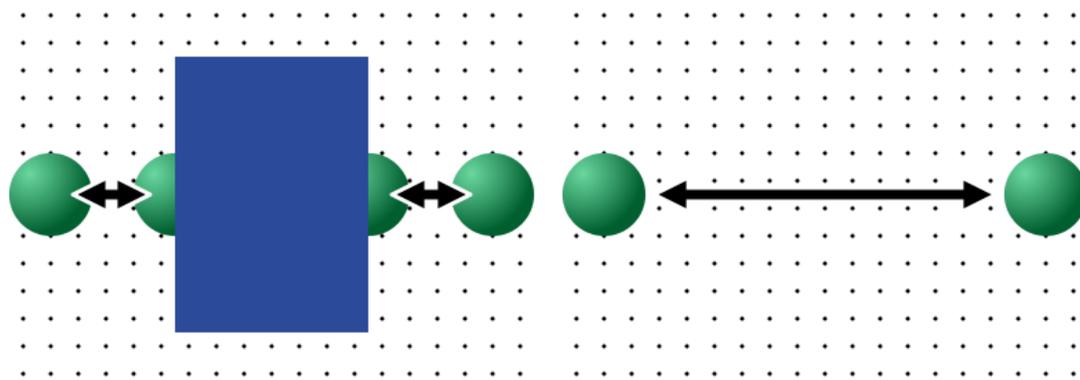
The interpretation of experiment 4 demonstrates the issues associated with individuating and identifying multiple occluded objects within the scene. The evidence from an object individuation task suggests that limited information is available for recall from memory. A finding that supports the limited nature of the naïve model. The fourth perceptual mechanism is discussed as a historical reference for an unexpected event. This is in contrast to the sensorimotor association which is a historical reference for a learned and therefore expected event.

Finally, a general discussion justifies the naïve model through a comparison with other similar research previously identified in the literature review. The interpretation of a fifth experiment uses the naïve model developed in the first four interpretations and applies it to a difficult object permanence problem; one that involves a stationary occluded object. The interpretation attempts to explain the empirical data using the naïve perceptual

mechanisms and without reference to object permanence. A discussion concludes that further development of the model is required in order to explain the empirical data associated with this study.

## 4.2 Experiment 1 – Spatiotemporal Completion (Eye Tracking)

### 4.2.1 Summary of Original Experimental Procedure



*Figure 3: Spatiotemporal Completion (Eye Tracking). Left: Occluded trajectory test display. Right: Continuous trajectory training display. Adapted from Johnson, Amsó and Slemmer (2003).*

This experiment is designed to test whether an infant can learn that a ball continues to exist while behind a box. If so, increased anticipatory looking (to the opposite side of the box from where occlusion occurs) should be shown while the ball is occluded. The experiment involves a ball moving along a straight path, and a box that acts as an occluder of the ball. The occluded trajectory display consists of a ball moving from one side of the display to the other behind the box, as depicted on the left of figure 3. The ball reverses direction when it reaches the end of the path and returns to the point from which it began. The process repeats until the end of the trial. The four balls visible in the figure are for illustrative purposes only and are intended to show the path of the ball behind the occluder; the experiment consists of only one ball. Eye movement was tracked during the trial to identify exactly where the infant was looking in relation to the display.

A second group of infants were initially presented with the continuous trajectory display, as depicted on the right of figure 3, and then with the occluded trajectory display. Both displays are identical except that the occluder is removed resulting in the ball moving along a continuous trajectory from one side of the display to the other, as depicted on the right of figure 3. Infants were presented with four, thirty-second continuous trajectory

training trials. Following the training trials, infants were presented with the occluded trajectory display. Eye tracking was used to determine differences in looking patterns before and after the training trials.

#### 4.2.2 Results and Original Interpretation

Infants of four and six months of age were tested with the occluded trajectory display in two separate groups. The first group was given training with the continuous trajectory display while the second was not. Four-month-old infants showed a significant increase in the amount of anticipatory looking at the opposite side of the box (from where occlusion occurred) after training with the continuous trajectory. The eye movement during training (continuous trajectory display) was consistently found to trail the object while the eye movement during testing (occluded trajectory display) was consistently found to lead the object. The researchers concluded that a limited amount of training provided sufficient information to allow four-month-old infants to learn object continuity and therefore predict object trajectory while partly occluded. The anticipatory eye movements of six-month-old infants did not improve with training which was interpreted as being because at this age infants can already generalise object continuity onto a new stimulus and therefore training is unnecessary.

Age	Original Interpretation of Result
4 Months	Anticipatory eye movements improve after training with the continuous trajectory display. Therefore, the infant is able to learn that an object on a trajectory behind an occluder continues to exist if given sufficient training.
6 Months	Makes anticipatory eye movements prior to training with the continuous trajectory display and does not improve after training. Therefore, the infant understands that an object on a trajectory behind an occluder continues to exist while occluded.

*Table 2: Summary of original interpretations of results for experiment 1*

#### 4.2.3 Naïve Interpretation

At six months of age the empirical results do not change between the two test conditions. It can be assumed that the original interpretation is correct and that six-month-old infants

have already learned rules concerning occluded trajectory continuity. Therefore, it is unlikely any new insight can be made through a naïve interpretation. Instead the focus of this interpretation will be on the change demonstrated by four-month-old infants when provided with training.

Prior to any experience with this display, eye tracking results showed that four-month-old infants do not successfully anticipate the path of the ball behind the box. Anticipatory eye tracking also did not improve with repetition suggesting that without additional information the four-month-old infant will remain uncertain as to the position of the ball's re-emergence.

In the naïve interpretation of these results it is assumed that the scene is perceived literally. Each iteration of the path consists of two instances of the ball appearing and then disappearing from existence. In this experiment there is no perception of depth from interposition and so all surfaces exist at the same level; nothing exists behind the box. There are also no learned rules that disallow the spontaneous disappearance of a surface; therefore, while the occlusion event is not understood, it is also not a violation of expectation. This might account for the fact that anticipatory eye tracking wasn't shown to improve with repetition; the disappearance of the ball is not an unexpected event given current perceptual knowledge and is accepted as normal behaviour. With repetition over time, a pattern consisting of the ball disappearing is learned and reinforced. The reason that the two trajectories are not initially perceived as two parts of a whole could also be related. Without an understanding of trajectory, the ball is just as likely to reappear from the point of disappearance.

After the continuous trajectory training phase, four-month-old infants showed significantly more anticipatory eye movement towards the far side of the box. The original interpretation suggests that with training, infants learned object continuity and created a representation of the trajectory behind the box. The implicit assumption is that prior to training, the infant either already knew that a background surface continued to exist behind the box or that this was also learned during the short training trials.

The differences in anticipatory eye tracking before and after training does suggest a recorded history of events. Without creating this history for later comparison, it would be

difficult to account for the difference in the findings before and after training. However, a mental representation of the continuous trajectory of the ball is an unnecessarily strong interpretation in this case. The comparison of this history to the sensory input from the occluded trajectory display does not necessarily require knowledge of occluded trajectory continuity. Rather, both conditions are still perceived literally, and the differences in empirical data are the result of sensorimotor associations made during training.

When perceived literally, the continuous trajectory training display differs from the occluded trajectory display in that movement never stops. There is no break in activity in which the infant has reason to perceptually split the path in two. This is not to claim that infants remain continually fixated on the ball for the full thirty seconds of each trial. However, the result is that visual scanning operates on segments across the entire width of the display, an area much larger than either of the two continuous trajectories observed in the occluded trajectory condition. At the end of the training phase, an association exists that causes scanning of the entire width of the display, and this is carried forward to the occluded trajectory condition. This could be the explanation for the improvements in the anticipation of disocclusion. One does not need to assume that the infant understands that the ball continues on its path behind the box. Instead, upon the ball's disappearance the infant continues scanning across the box because of a learned motor action. The novelty of the disappearance event that caused the infant to stop scanning in the pre-training condition is overridden by the sensorimotor association and scanning continues until the ball re-emerges from the other side of the box. The two trajectory parts are now observed together and repetition only strengthens the sensorimotor association of scanning across the box.

This doesn't explain how, without object knowledge, the infant is able to associate the procedural memory of scanning across the display with the movement of the ball. The answer is that rather than an association between the ball itself and the motor action, the association is between positional change in general and the motor action. The infant doesn't need to have a representation that is specifically bound to the featural information of the ball because visual scanning across the display is now the default action for any positional change.

Positional change is the lowest level information available to the perceiver when a surface (a region of the view) changes position in space. Prior to any object knowledge the only aspect of motion that is detected by the perceiver is a surface change from position A at time  $t_1$  to position B at time  $t_2$ . In this sense, the ball moving along a continuous trajectory becomes a surface changing position in space over time. Although at first this sounds like semantics, it is important to appreciate that a trajectory implies an understanding of continuity, velocity, and direction, while a change in position does not. Changes in position over time, and the inherent reward that comes from tracking them successfully, are responsible for the infant's initial interest in a stimulus and the visual scanning that follows.

#### **4.2.4 Discussion and Justification**

This naïve interpretation highlights a common, but often ignored interpretive bias of developmental researchers. The bias is caused by the macro scale at which infants preferential looking is often judged. In most studies this is binary and the infant either is, or is not, looking at the display. Researchers often use this binary judgement to infer that infants gather enough information to perceive the scene, and that therefore, the problems they demonstrate are based only on a misunderstanding of that information. This experiment is one of the few available in which eye tracking is used to empirically demonstrate the differences in patterns of visual scanning, and consequently the information that arrives as sensory input.

The original interpretation of the results of this experiment is that infants can learn object trajectory continuity with a small amount of training. This assumes object knowledge including continuity, segregation, and re-identification. The naïve interpretation argues that object knowledge is not required to explain the results of this experiment. Instead, a combination of the reward inherent to tracking positional change and the procedural motor memory that emerges from repeated action provide a plausible explanation for the empirical data.

The reward inherent to tracking positional change increases and decreases based on novelty. The novelty of the positional change decreases if it is easily anticipated because it has been learned through repetition. Visual scanning reduces the visible positional

change by keeping the target object in the centre of view, and results in decreased reward and a strengthened procedural memory.

The intention of this interpretation is not to reject the inclusion of featural information in a memory of events, rather it is to show that featural information need not automatically be bound to the event from the beginning. An episode of events initially consists only of spatiotemporal motor action information. This might be because the surface is moving too quickly to easily identify or because the trajectory is novel. As more sensory data is collected, featural information (the image of the ball) is bound to the spatiotemporal information (the trajectory of the ball) creating a sensorimotor association.

This is also sufficient to explain the effects of habituation on the familiarisation trials of this experiment. The first trial is novel to the infant because the positional change of surfaces within the scene is new. With repetition the reward provided by tracking the positional change decreases as the procedural motor memory strengthens. Essentially, tracking the positional change becomes automatic and uninteresting. Habituation to a stimulus is the lack of reward associated with predicting positional change.

## 4.3 Experiment 2 - Spatial Completion (Including Eye Tracking)

### 4.3.1 Summary of Original Experimental Procedure

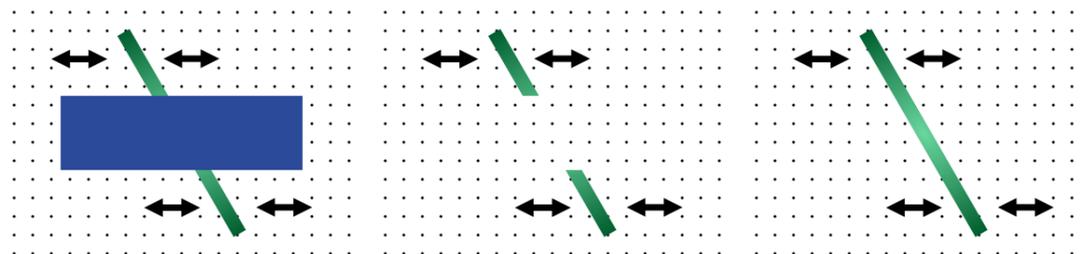


Figure 4: Spatial Completion (including eye tracking). Left: Occluded rod habituation display. Centre: Broken rod test display. Right: Continuous rod test display. Adapted from Johnson and Náñez, (1995).

The rod-and-box experiment is designed to test an infant's ability to use motion to represent a partially occluded surface as a complete entity. The infant is first habituated to a rod translating horizontally behind a box. After habituation to this stimulus, the infant is presented with two successive test displays that involve the same pattern of horizontal rod motion, but do not contain the occluding box. The first test display consists of a broken rod and the second consists of a complete rod. The order of the test displays is counterbalanced across the infants participating in each trial. The researchers contend that if the infant looks longer at the complete rod test display, it means that the broken rod is familiar. This in turn means that the infant perceives the rod as two distinct parts. If the infant looks longer at the broken rod test display, then the opposite is true: the infant perceives the rod as an individual object, and therefore, spatial completion of the rod must have been achieved. In a separate study, eye tracking was used to determine differences in the looking patterns of perceivers and non-perceivers of spatial completion (Johnson, Slemmer & Amso, 2004).

### 4.3.2 Results and Original Interpretation

This experiment originally tested four-month-old infants using a wooden rod-and-box display (Kellman & Spelke, 1983). After habituation, infants looked longer at the broken rod which was inferred as being because they could perceive the complete rod. These results were some of the first to conflict with Piaget's sensorimotor theory that perception

requires action, and were interpreted as possibly being due to innate knowledge of certain properties of an object. Johnson and Nanez (1995) repeated the experiment using a computer-generated stimulus in order to eliminate any cues to depth. They found that during the test phase four-month-old infants looked longer at the broken than at the complete rod, while two-month-old infants did not look preferentially at either test display. This was interpreted as demonstrating that the perception of spatial completion is developing at around two months of age, and by four months of age, infants are capable of spatially completing a surface without depth information. They speculated that the perception of concurrent motion is an important factor for spatial completion in infancy but concluded that there was not yet enough evidence to speculate further.

Johnson et al., (2004) found significantly different patterns of looking in the rod-and-box condition, between so called perceivers and non-perceivers of spatial completion. Infants who looked preferentially at the broken rod test display also showed increased visual scanning with the trajectory of the rod parts, within an individual rod part, and across the two rod parts. The authors of the study decided not to postulate whether the difference in observation was the cause, or the effect of the perception of spatial completion.

<b>Age</b>	<b>Original Interpretation of Result</b>
All (control condition)	Infants look equally at both, therefore, neither condition is inherently novel.
2 Months	Infants look equally at both, therefore, spatial completion must be developing at this age.
4 Months	Infants look longer at broken rod, therefore, complete rod is familiar.

*Table 3: Summary of original interpretations of results for experiment 2*

Age	Result
Non-Perceivers (2 Months)	Increased scanning of regions of the display that are likely less relevant to the perception of spatial completion.  Decreased scanning across the two rod parts.
Perceivers (4 Months)	Increased horizontal scanning with the motion of the rod part.  Increased within rod part scanning.  Increased scanning across the two rod parts.

*Table 4: Summary of original interpretations of results for experiment 2 (eye tracking)*

### 4.3.3 Naïve Interpretation

Two-month-old infants who do not look preferentially at the broken rod test display, demonstrate less visual scanning of the important regions of the habituation display. They spend less time following the horizontal trajectory of the individual rod parts than their older counterparts. Given that the positional change associated with the habituation display consists entirely of horizontal translation, it seems that tracking the rod in its entirety is a learned skill.

It is possible that because of the patterns of visual scanning, only enough sensory data is collected to partially perceive a rod part and its trajectory. Essentially, many short episodes of events are perceived, but because there is so much positional change within the scene no motor action is repeated often enough to demand addition into procedural memory. If the two rod parts are only partially perceived, it is not surprising that after habituation neither test display is more novel than the other; the two-month-old infant ends up with a partial memory of the display that subsequently matches both test displays.

Two-month-old infants are more interested in exploring a range of different and discrete positional changes within the display than they are with scanning the length of the rod as if it were an individual object. This could be attributed to the amount of positional change occurring in the rod-and-box display at any one time. The ball from experiment 1 can be described by a single spatial position change over time. In contrast, the rod in this experiment consists of at least two different spatial positions. The bottom rod part is qualitatively different from the top, both because of the distance separating the two, and

because of their different relationships with the box. Visual scanning must jump around regions of the display in order to perceive the scene, as opposed to just following one position, as was assumed to be the case in experiment 1.

To perceive the trajectories of the two rod parts as cohesive, the infant must progress from using a seemingly unplanned pattern of visual scanning to a more controlled pattern that primarily involves the length of the rod. The infant must first learn to track the horizontal change of the rod at different points in order to build up a series of different episodes; at least one for the motion associated with the top and another for the motion associated with the bottom part of the rod.

Disparate regions of an object can only be associated after their trajectories are defined individually. As was discussed in the interpretation of experiment 1, the individuality of a surface on a trajectory will persist for the duration of the time it is being directly tracked by the eye. The spatiotemporal information defining the trajectory accumulates while visual attention is directed toward following the surface and stops when attention is redirected elsewhere. Essentially, focusing on and following a surface makes it remain the same individual over time. This is true for both the top and the bottom of the rod. Following the trajectory of either as it translates horizontally is recorded as a discrete episode of time and includes the spatiotemporal information associated with the trajectory. As soon as visual attention is directed elsewhere, a new episode is created that is isolated from the others.

The second perceptual mechanism of the naïve model is used to describe the association of the individual rod parts. Following a surface on a repetitive trajectory causes the novelty of positional change to return to near baseline level (the level it is prior to any positional change) because the visual scanning becomes an automatic response. This provides the opportunity to notice similar positional change that may be occurring elsewhere in the display. With experience this results in the ability to switch between the two trajectories, and is evidenced by the increased scanning within and across rod parts as demonstrated by the infants originally interpreted to perceive spatial completion.

The infant is eventually habituated to the display including the rod parts and the box. During the test phase, both the complete and broken rod test displays are novel because

the box is now absent. The infant isn't necessarily aware of what has changed but the sensory data and the history of events are inconsistent. When the infant scans along the rod part and finds the gap, the missing positional change increases attention. This may seem counter-intuitive because the gap was already apparent during habituation; however, because the box is removed a new set of episodes are created and the sensorimotor associations need to be created again. In contrast to the broken rod test display, scanning from top to bottom along the complete rod test display contains no gap, removing the inconsistency in positional change. This time only one episode is required to represent the complete length of the rod.

Unfortunately this claim cannot be supported by empirical data because scanning patterns during the test conditions were either not collected or not published. Nonetheless, this remains an interpretation of results that doesn't rely on the mental representation of the complete rod existing behind the box.

#### **4.3.4 Discussion and Justification**

The original interpretation is that younger infants do not perceive spatial completion and older infants do, and that the eye tracking data suggests that either this is caused by, or is the effect of visual scanning (Johnson, Slemmer & Amso, 2004). Although a plausible interpretation of the data, this offers little explanation as to the development between age groups. What is it that changes patterns of looking between four and six months of age? The naïve interpretation differs from the original in that it offers an explanation for the difference in findings between the two age groups; i.e. an explanation that rejects the notion of spatial completion in exchange for an alternative that only requires the ability to track the horizontal trajectory of multiple surfaces. This explanation allows multiple surfaces that share a common motion to be perceived as a single object without needing to assume that they are physically connected.<sup>2</sup> It is important to point out that this alternative requires a fundamentally different interpretation of the test phase of the experiment. The original experimental procedure relies on one test display being novel and the other being familiar. Neither is entirely true in the naïve interpretation. Instead,

---

<sup>2</sup> This is a similar effect and perhaps the mechanism behind the perception of a point-light walker as a coherent object. See section 2.3.2.

both displays are novel, but one less so than the other. This difference in data analysis will be seen frequently in the following interpretations.

Whether or not the naïve interpretation is a more plausible explanation of this experiment, the intention is to point out that infants of different ages are not collecting the same sensory information and that learning to do so competently is a prerequisite for associating regions of a surface into an individual object. It is problematic to assume that the trajectory of either rod part has any intrinsic meaning to an infant. Just because the researcher perceives the display as two rod parts sharing a common trajectory doesn't mean that this sensory information is also available to the infant.

The naïve interpretation is supported by more recent evidence in which the experiment was repeated with newborn infants using stroboscopic rather than horizontal translatory motion (Valenza, Leo, Gava & Simion, 2006; Valenza & Bulf, 2011). Stroboscopic motion involves spatiotemporal oscillation of the stimulus so that it appears to jump between two spatial locations over time. Young infants show a preference for flicker over other types of motion and a flicker detection system almost as good as that of an adult by twelve weeks of age (Regal, 1981).

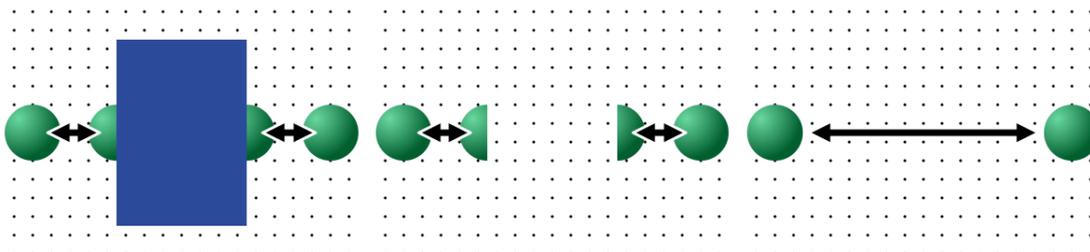
When stroboscopic motion is applied to the rod-and-box experiment the resulting stimulus consists of an occluder and a broken rod that jumps between two spatial locations. The rod remains stationary for approximately half a second between oscillations. Otherwise the experimental procedure is equivalent to the original experiment. The infant is habituated to the rod moving behind the occluding box and then tested successively with complete and broken rods undergoing the same stroboscopic motion but minus the occluding box.

The results of the rod-and-box experiment with stroboscopic motion conflict with those of the original experiment. After habituation, the newborn infant looks longer at the broken rod test display which is interpreted as the apparent ability to spatially complete the broken rod. The authors propose that spatial completion is essentially an innate concept that is obscured at birth only because the immature visual system is unable to successfully track a laterally translating stimulus (Valenza et al., 2006). The naïve interpretation is that this conclusion is technically accurate, although the ability to actually

spatially complete the rod at this age remains a point of contention. The stroboscopic motion eliminates the need to learn the rod's horizontal trajectory, and because both rod parts essentially disappear and re-appear together, it is more likely that they will both be part of the same episode.

## 4.4 Experiment 3 – Spatiotemporal Completion

### 4.4.1 Summary of Original Experimental Procedure



*Figure 5: Spatiotemporal Completion. Left: Occluded trajectory habituation display. Centre: Discontinuous trajectory test display. Right: Continuous trajectory test display. Adapted from Johnson, Bremner, Slater, Mason, Foster and Cheshire (2003).*

This experiment is the precursor to experiment 1, which also tests spatiotemporal completion but using a procedure more suitable to collecting eye tracking data. This experiment begins by presenting infants with the ball moving along a partially occluded trajectory from one side of the display to the other. The ball is hidden as it moves behind the box. The ball reverses direction when it reaches the end of the path and returns to the point from which it started. This repeats until the end of the trial. After the infant is habituated to the occluded trajectory display the test trials begin. The occluding box is not present in either test display. The continuous trajectory test display consists of the ball on an unobstructed trajectory from one side of the display to the other. The discontinuous trajectory test display consists of the ball moving along the same trajectory. The difference being that the ball disappears and reappears when it reaches the previous position of the box, as it would if the box was present. To an adult, the ball appears to become occluded by the background surface, an impossible event.

This experiment is designed to test an infant's ability to represent an object as continuous as it moves in a straight line behind an occluder. This is calculated by recording the time spent looking at the two test displays. The researchers assume that if the infant looks longer at the continuous trajectory test display it means that the discontinuous trajectory is familiar. They conclude from this that the infant perceives the ball as disappearing from existence while occluded. If the infant looks longer at the discontinuous trajectory test display then the opposite is true and object trajectory continuity is being perceived.

#### 4.4.2 Results and Original Interpretation

Infants of four and six months of age were tested. After habituation, four-month-old infants looked preferentially at the continuous trajectory test display while six-month-old infants looked preferentially at the discontinuous trajectory test display. The researchers interpret the results as showing that six-month-old infants perceive the ball as continuous while occluded behind the box. They are aware that the ball continues to exist during the motionless occlusion event, and expect its reappearance. The implicit assumption is that the infant should understand that a foreground surface, in this case the box, must be occluding the surface that surrounds it. There is a difference between this and the assumption that the ball is occluding the background because the ball is changing position and therefore revealing the background surface as it moves. Instead, the box remains stationary; the implication being that the perception of any surface existing behind it must be a representation based on pre-existing perceptual assumptions. The original interpretation is that four-month-old infants do not understand that the ball continues to exist while behind the box. The researchers conclude that some mechanism for perceiving occluded trajectory continuity develops between these ages.

<b>Age</b>	<b>Interpretation of Result</b>
All (control condition)	Infants look equally at both.
4 Months	Infants look longer at the continuous trajectory because it is novel. Therefore, the infant does not understand that an object on a trajectory behind an occluder continues to exist.
6 Months	Infants look longer at the discontinuous trajectory because it is novel. Therefore, the infant understands that an object on a trajectory behind an occluder continues to exist.

*Table 5: Summary of original interpretations of results for experiment 3*

#### 4.4.3 Naïve Interpretation

As discussed in the naïve interpretation of experiment 1, four-month-old infants perceive the occluded trajectory display literally. For this interpretation, it is assumed that four-month-old infants have not yet acquired the ability to segregate objects. Therefore, surfaces including the ball can appear and disappear without violating any pre-existing object knowledge. The only novelty in these displays is the result of positional change.

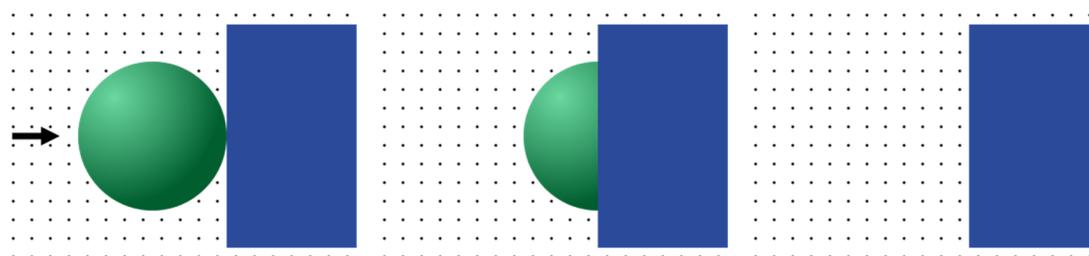
As a result, the box is an unimportant feature of the display as it doesn't change position making the discontinuous trajectory test display very similar to the habituation display. The continuous trajectory test display is most interesting to four-month-old infants because when perceived literally it has the most differences compared to the habituation display.

In the original interpretation six-month-old infants look longer at the discontinuous trajectory test display after habituation because this is novel to them. It is claimed that the continuous trajectory test display is familiar because this is what was actually being perceived during habituation. This interpretation makes implicit assumptions about object knowledge including that the infant understands the concept of interposition (the perception of depth based on partial occlusion). Interposition in turn depends on an understanding of occlusion, solidity (two objects cannot simultaneously occupy the same position in space), and a feature of trajectory continuity (objects do not spontaneously disappear from existence). Each of these concepts requires an explanation as to its origin.

The naïve interpretation relies on the difference between spatial changes to explain these concepts in regards to the current empirical data. At a low level, surface segregation is the difference between two spatial positions over time. Two pieces of information important for segregation emerge as the ball moves from position A to position B. Firstly, the ball is now at position B, but of equal importance is that the ball is also no longer at position A and the background surface is now revealed. Without this second piece of information all that can be perceived is positional change on a single layer of depth; the background surface could conceivably be changing position with the ball. Essentially, perceiving this swapping of surfaces is the low level mechanism of segregation. The surface segregation mechanism operates by observing change over time and therefore requires motion.

Recording both the current and previous positions of an object is not an innate perceptual ability but is instead learned with training. This is why the ball moving across the background is not necessarily perceived as two individual surfaces. Visual scanning follows the target object, in this case the ball. The background is largely ignored because there is less interest in targeting this motionless surface.

So how does a four-month-old infant acquire the segregation mechanism available to a six-month-old while only capable of detecting positional change? Repeated exposure to the stimulus is clearly not the answer or habituation to the occluded trajectory display would be sufficient. At best, repetition only reinforces the existing understanding, which in the case of four-month-old infants, is that the ball is disappearing from existence. Instead, as discussed in experiment 1, a different perspective requires a similar but new stimuli. In experiment 1 the eye tracking of the four-month-old infant in the ball-and-box condition changed as a result of pre-exposure to the continuous trajectory condition. It was argued that these findings didn't require the original interpretation which assumed that the infant learned occluded trajectory continuity. The evidence demonstrates that the stimulus was observed differently as a result of sensory input competing with a prior memory.

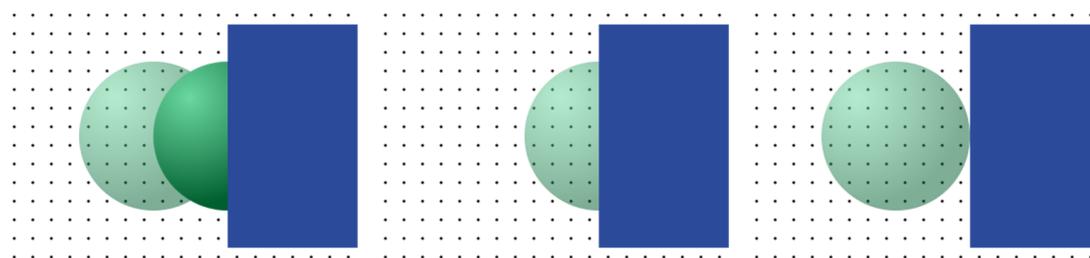


*Figure 6: Occlusion Event. Left: Episode of ball prior to occlusion. Centre: Episode of partial occlusion of ball. Right: Episode of full occlusion of ball.*

In this experiment, the competing sensory information is a result of building relationships between multiple spatial positions on the ball. If the ball is only perceived as a single spatial position change over time, then the infant will only ever have reason to perceive the change at the new position. In this case, the perception of multiple spatial positions occurs as a result of the occlusion of the ball. Assuming that the infant is capable of perceiving the occlusion event in sufficient spatiotemporal detail, when the ball moves behind the box it is possible to observe different amounts of its surface at different times. If the perceptual capabilities of the infant are insufficient and partial occlusion is not perceived, the infant cannot succeed at object segregation using this mechanism.

Fortunately, the disappearance and reappearance of the ball actually serve to increase the likelihood of perceiving partial occlusion. As evidenced with stroboscopic motion in the

discussion of experiment 2, the transformation from no positional change within the scene to the appearance of an object, or vice versa, is highly salient and attention is immediately redirected to the source of the positional change. If infants are observing the partial occlusion event literally, then the partially occluded ball is perceived as half a ball. The other half of the ball ceases to exist. Observing half of a ball is problematic to the continuity of the ball on its trajectory because half of a ball is an entirely new instance of the surface and is distinct from the rest of the scene. Figure 6 helps to demonstrate more clearly why the partial occlusion event is important. If the infant only perceives the left and right frames of figure 6, the position at which the ball disappears during occlusion immediately becomes the end of the episode. There is no reason to perceive the ball as interacting with the box. The effect of this is evidenced in the empirical data by the four-month-old infants' apparent belief that the ball has disappeared from existence.



*Figure 7: Segregation Mechanism (partially transparent balls indicate previous episode). Left: Complete ball compared to half occluded ball. Centre: Half occluded ball compared to full occlusion. Right: Complete ball compared to full occlusion.*

The segregation mechanism requires that two positions are observed during an occlusion or disocclusion event. Both the leading and trailing edges of the ball must be represented separately, but remain part of the same object. As discussed in the interpretation of experiment 2, this can occur given sufficient experience of the two positions when on the same trajectory and with the same velocity, factors that are certainly apparent with this stimulus. However, unlike the top and bottom rod parts from experiment 2, the leading and trailing edges of the ball are only represented separately when the partial occlusion event is perceived. The occlusion event can be interpreted as involving three episodes of time, as depicted in figure 6. The first episode consists of the ball on its trajectory toward the box and ends when the ball is adjacent to the box because of the featural change. The

second episode consisting of half a ball is created and exists briefly until the ball is fully occluded. The third episode begins when the ball is fully occluded by the box. Because the featural information of the half ball is novel but appears on the same spatiotemporal trajectory, it is compared with the featural information of the previous episode, a complete ball, as depicted on the left of figure 7. A similar comparison occurs between the second and third episodes, as depicted in the centre of figure 7.

This process is essentially the operation of the segregation mechanism described earlier; the previous image is compared to the current image because of novel differences in similar episodes of events. This comparison marks a point of surface interaction in the representation and includes the three episodes including the ball, the half ball, and the box. As a contrast it is also beneficial to this argument to describe the comparison of episodes when partial occlusion is not observed. If for example the ball is moving so quickly that it disappears without partial occlusion being observed, then the comparison is between a complete ball and full occlusion of the ball, as depicted on the right of figure 7. This may provide some segregation of the ball and the background, but does not provide segregation of the ball and the box.

This mechanism of object segregation is supported by the evidence of visual scanning from experiment 1. Infants in the continuous trajectory training condition were shown to consistently trail the ball while infants in the occluded trajectory test condition were shown to consistently lead the ball. The authors couldn't determine a concrete explanation and contended that the reason was likely physiological rather than psychological (Johnson et al., 2004). Rather than being entirely physiological, these results could indicate that infants are only perceiving positional change on a single layer of depth when there is no clear occlusion event, as is the case with the continuous trajectory training condition. By trailing the ball infants do not have the opportunity to process the differences between layers. In contrast, scanning that leads the ball suggests that the infant knows how to deal with a partial occlusion event. By leading the ball infants can process both the complete ball before occlusion and the trailing edge of the ball as it is occluded.

At this point it is important to reiterate that the naïve interpretation of this experiment does not provide the infant with the ability to mentally represent the ball as it moves

behind the box. Instead, all that is learned is the segregation of the surfaces; in this case, that the ball, the box, and the background are not regions of the same surface. An understanding of surface segregation is a prerequisite for a variety of other important object concepts including object continuity and object solidity. Both of these concepts are only slight extensions of the surface segregation mechanism.

Object continuity – the understanding that an object shouldn't spontaneously disappear from existence – is a learned skill, as is evidenced by the empirical data from this experiment. A control group including both four and six-month-old infants who were not habituated to the occluded trajectory showed no preference for either the discontinuous or the continuous trajectory test conditions. This indicates that a ball disappearing without an occluder is not an intrinsically novel stimulus. The segregation mechanism described allows an infant to learn that a representation of an occlusion event should involve the three episodes described earlier as zero, partial, and full occlusion of the ball.

After habituation to the occluded trajectory, six-month-old infants looked preferentially at the discontinuous trajectory test condition. The original interpretation is that this is because the infant is mentally representing the continuous trajectory and therefore it is most familiar. The naïve interpretation does not support this claim because the infant has not learned that the trajectory is capable of continuing behind the edge of the box. Instead, the naïve interpretation of this result is that the occlusion rule learned during habituation is inconsistent with the arriving sensory information of a ball being occluded by a background surface. Both test conditions are novel due to the absence of the box alone, but the continuous trajectory condition is not also inconsistent with the newly learned occlusion event, and is therefore less novel.

This explanation of the empirical data requires a third type of perceptual mechanism be added to the naïve model. This type is encoded based on an inconsistency between sensory information and learned object knowledge. In this case, the object knowledge is related to occlusion and object segregation. The six-month-old infant knows that the occlusion event involves episodes consisting of the occluding box because these events were the most novel during habituation, and as a result were encoded into the representation. Observing

the same occlusion event but without the box is novel because the absence of the box requires a new representation that is inconsistent with the original.

#### **4.4.4 Discussion and Justification**

The naïve interpretations of experiments 1 and 2 begin to describe motion at a low level, but largely ignore the interaction between the different surfaces because to explain the empirical data in those cases requires only a single layer of depth. This interpretation continues from that foundation and describes a perceptual mechanism for the interaction between background and foreground layers using only information available as a result of spatial position change.

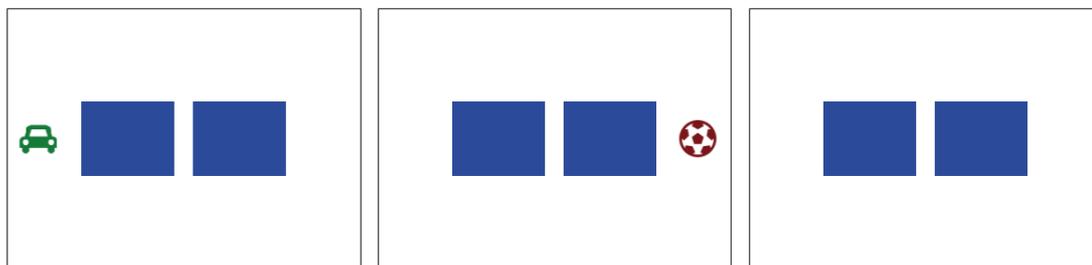
Object segregation is often assumed to be an innate ability of infants. Infants are able to grab and hold objects in the real world which would imply that they are being individuated from the background. This may be true, however, there is a difference between holding and perceiving an object in three-dimensional space, and seeing an object in the distance and segregating it from the background. All of the experiments so far have been composed of two-dimensional images rendered on a computer. Depth perception by biological means should be irrelevant, in which case object segregation requires explanation. In contrast to the original interpretation of results, the naïve interpretation offers an explanation for the developmental differences between the two age groups tested. The younger infants are unable to segregate the ball and box because they do not perceive the stimulus in a sufficient amount of spatiotemporal detail, and therefore have not learned to associate episodes of zero, partial, and full occlusion into a cumulative episode of events.

A simple modification to the original experimental procedure might be all that is required to confirm or deny this naïve interpretation. The habituation, discontinuous, and continuous displays all remain the same except that the continuous motion of the ball is replaced with apparent motion. Apparent motion originated as a definition for film and is implemented as frames on a film strip. The apparent motion in a film appears as a continuous stream of live action to the observer; however, the apparent motion I propose for this experiment would be so slow that each frame is easily individuated. In one series of experiments the ball would follow its trajectory one frame at a time until it was adjacent to the box and would then disappear. In another series of experiments the ball would

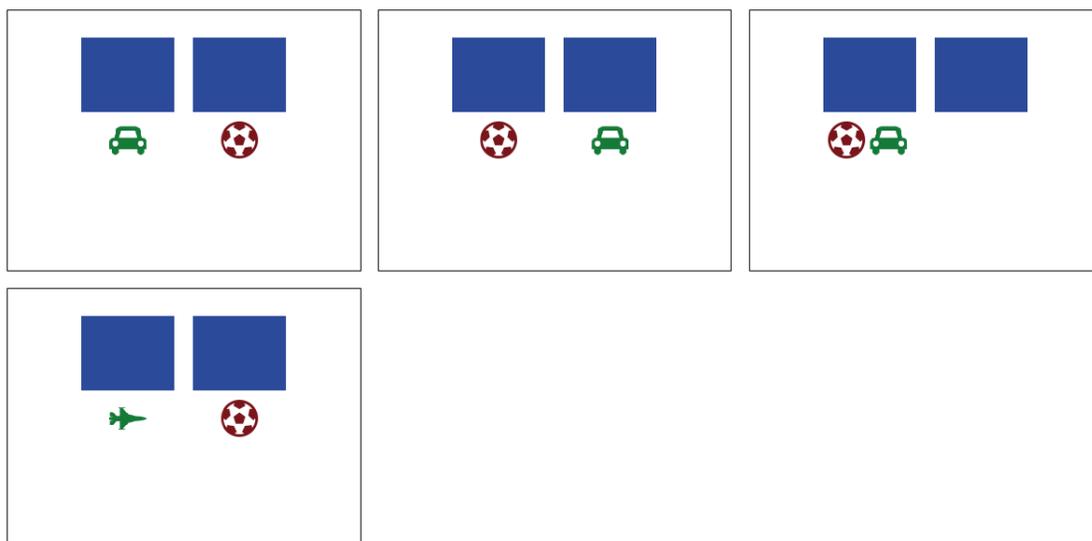
follow its trajectory but the ball would be partially occluded by the box for one frame before disappearing. If there are no differences between these results and those of the original experiment then there is additional support for a theory of occluded trajectory continuity as it was originally interpreted. However, if infants look preferentially at the discontinuous trajectory when habituated to a partial occlusion event, but look preferentially at the continuous trajectory when habituated to the ball disappearing without partial occlusion, it can be contended that object segregation is unavailable until the infant is provided with sufficient training.

## 4.5 Experiment 4 – Object Individuation and Identity

### 4.5.1 Summary of Original Experimental Procedure



*Figure 8: Object Individuation Familiarisation Events. Left: First object emerges from left occluder, second object is occluded. Centre: Second object emerges from right occluder, first object is occluded. Right: End of familiarisation trial, both objects occluded. Adapted from Mareschal and Johnson (2003).*



*Figure 9: Object Individuation Test Events. Top Left: Objects revealed in expected configuration. Top Centre: Objects revealed, location of objects is switched. Top Right: Objects revealed, location of objects has switched and both appear behind one occluder. Bottom Left: Objects revealed, one object is replaced with an unfamiliar object. Adapted from Mareschal and Johnson (2003).*

This experiment is designed to determine the features of an object that are maintained during occlusion. Prior to testing, infants are familiarised with two objects as depicted in figure 8. The familiarisation phase begins with the two objects hidden behind two occluders. The first object emerges from behind one of the occluders, remains stationary for five seconds and moves back behind. The second object then emerges from behind the

other occluder, remains stationary for five seconds and then moves back behind. This repeats another four times before the familiarisation phase is complete. Infants are pre-allocated into groups that each correspond to one of the test conditions depicted in figure 9. After a five-second pause where only the occluders are visible, they are raised to the top of the display and reveal the objects as per one of the test conditions. The interpretation that infants as a group experience a violation of expectation is based on the average looking time for the test condition. The baseline test condition, as depicted on the top left of figure 9, is used as the control condition to which other looking time results are compared. Looking times above the control condition imply a violation of expectation.

#### **4.5.2 Results and Original Interpretation**

The objective of this study was to determine the variances in the individuation and identity information maintained during occlusion for different types of stimuli. Trials involved three types of stimuli: female faces, asterisks, and manipulable toys. The researchers found different patterns of preferential looking depending on the type of stimuli observed. When the stimuli consisted of either female faces or asterisks, infants showed increased looking for a test display in which the features of one of the objects had been altered (a face had been swapped for a new face, or an asterisk had changed colour), as depicted on the bottom left of figure 9. In contrast, when the stimuli consisted of images of manipulable toys, infants showed increased looking only for the test display in which both objects appeared behind one occluder, as depicted on the top right of figure 9. For all stimuli tested, infants showed equivalent looking times for both the control and the object reversal conditions, as depicted on the top left and top centre of figure 9 respectively.

The researchers interpreted these results as demonstrating differences in the visual processing of different types of stimuli. Evidence suggests two separate visual object processing streams exist; one for spatiotemporal and another for featural information (Milner & Goodale, 1995). The results between categories seem contradictory. The asterisk and female face categories are outliers and challenge previous evidence that suggests spatiotemporal information is encoded ahead of featural information (Leslie, Xu, Tremoulet & Scholl, 1998; Xu & Carey, 1996).

Age	Interpretation of Result
4 Months (All Categories)	Similar amounts of looking for both test conditions that consist of one object behind each occluder. Therefore, spatiotemporal but not featural information is maintained.
4 Months (Manipulable Toys)	Infants looked preferentially at the test condition in which two objects appeared behind a single occluder. Therefore, spatiotemporal information is more important.
4 Months (Asterisks, Female Faces)	Infants looked preferentially at the test condition in which the features of one object changed. Therefore, featural information is more important.

*Table 6: Summary of original interpretations of results for experiment 4*

### 4.5.3 Naïve Interpretation

If the events of the experiment are split into shorter episodes of time, it becomes apparent that only a single instance of an object need exist during the familiarisation trials. Surfaces are only ever observed one at a time prior to the test trial, and because the naïve model does not support occluded trajectory continuity, the location of each surface's re-appearance is irrelevant. As the two trajectories are never observed simultaneously, the only reason to perceive them as two instances would be to divide them by feature. The contradictory results between categories suggests that there may be one instance of an object bound to multiple instances of featural information.

The introduction of the objects in the familiarisation trials is similar to the occlusion and disocclusion of the ball behind the box in the interpretation of experiment 3. Indeed, when either of these events are perceived literally they consist of two separate trajectories. The primary difference is that in this experiment there is a small gap between two occluders so that there are no cues to suggest that the two objects could actually be the same individual. As the naïve model currently doesn't allow for the representation of a surface on an occluded trajectory, this factor has no bearing on the interpretation.

If there is no perception of occluded trajectory continuity, it is likely that the point of occlusion would mark the spatial position encoded into the representation. In this case, the position at the edge of the occluder where the occluded object was last observed. Therefore, rather than expecting the object to appear behind the occluder as it does in the

test trials, it is more likely that the infant expects the object to appear from the edge of the occluder, as occurs during the familiarisation trials. Consequently, every test display is novel because the location of object re-appearance during the test trial is unexpected.

As the empirical data shows, swapping the location of the two objects does not increase preferential looking regardless of the category of stimuli. The original interpretation of this was that infants are able to maintain a spatiotemporal or a featural representation of an occluded object, but not both. Categories of stimuli differed in regards to featural properties. Objects from all categories were the same size, but the asterisk and female face categories also all shared a similar shape. The asterisks only differed by colour, and the shape of the female faces were all similar. In contrast, the manipulable toys were all different shapes and colours. The result that asterisks and faces can be identified by feature but toys cannot, suggests that the similarity in shape allows for a more detailed encoding of featural information.

The naïve interpretation is that during familiarisation the representation consists of the last observed surface trajectory. The spatiotemporal information for manipulable toys are encoded individually and so only the features of the last observed object are available. Asterisks and female faces are encoded differently. Each time the object is encoded, the associated featural information remains similar enough that it isn't entirely overwritten. The encoding can consist of two colours of asterisk for example. However, the added featural detail decreases the accuracy of the spatiotemporal information, and the position of the last occlusion event can be mistaken.

Infants in the manipulable toys category look preferentially at the test display consisting of two objects behind a single occluder because the proximity of the objects is novel when only one object is expected. Increased processing is required to identify the correct featural information. Observing two objects at a distance is also fairly novel, but not to the same extent. When the category is asterisks or female faces, the surface feature test display is most novel because observing two objects regardless of their position is more expected than an unfamiliar object.

#### 4.5.4 Discussion and Justification

The original interpretation of results makes the assumption that four-month-old infants maintain a representation of two individual objects during occlusion. However, the results for the three categories of stimuli are contradictory. The researchers contend that because the toys are manipulable there is an inferred affordance for action and so a representation would rely primarily on spatiotemporal information. In contrast, the objects in the other categories are not typically manipulable and so featural information is more important for representation. Essentially, the original interpretation is that depending on the category of stimuli either information related to spatial position, or information related to image features is encoded, but not both.

The naïve interpretation suggests that spatiotemporal information is primary and featural information is bound later, but that the binding can be rather ambiguous depending on similarities amongst different objects. This interpretation is similar to the original, but differs on the explanation for the contradictory results. The original interpretation requires that infants have categorised manipulable and non-manipulable objects, while the naïve interpretation requires the ability to compare surface differences. The naïve interpretation is supported by the evidence from experiment 1. It will be recalled that the representation of the trajectory of the ball was described as being primarily spatiotemporal with featural information bound over time. This sensorimotor association allowed for the continuation of visual scanning over temporary occlusions of the ball provided it remained on the same trajectory. The featural information is clearly not always available.

This interpretation also differs from the original in that there is no assumption that the two objects must be considered as individuals. The evidence suggests that the representation typically only contains useful information for the object that was last seen. This seems like a plausible extension of the findings of interpretation 2 in which infants seemed to build a partial representation of events over time, rather than instantly having a complete snapshot of the entire scene after the first view.

## **4.6 General Discussion of the Naïve Model**

Each of the four naïve interpretations included a brief discussion about the novelty of the bottom-up approach when compared to the original interpretation. These demonstrated the differences in assumptions about object knowledge that were necessary in each case, and in so doing, justified the use of a bottom-up interpretive approach.

The naïve model that emerged as a result of the collection of interpretations also requires analysis and justification. It is important to understand the broader implications of this model. As discussed in chapter 2, there is extant literature that empirically demonstrates the issues inherent to various experimental procedures. In contrast, the objective of this research was to model the development of early object knowledge using the bottom-up approach across a range of experiments. This section will present a summary of the model including the assumed knowledge, an interpretation of the perceptual primitives involved, and a comparison of these against the equivalent primitives/concepts from similar models. This section ends with a discussion on the future development of this model, and whether it could form the foundation of higher level conceptual knowledge.

### **4.6.1 Summary of the Model**

The set of naïve interpretations have contributed to the development of a theoretical model of very early object knowledge. The model describes the segregation of visible surfaces, the re-identification of trajectories, and allows for the differentiation of occlusion and disappearance events. Moreover, it does so without a dependence on innate abstract knowledge.

#### **Contribution of Interpretation 1**

The model assumes that the infant is innately able to detect visible movement, and that tracking this movement over time is intrinsically rewarding. This is best demonstrated in the interpretation of experiment 1 where evidence from eye tracking of a spatiotemporal completion task is used to argue that following a surface on a trajectory, even during partial occlusion requires no object knowledge. Instead, a combination of sensorimotor associations and an interest in positional change is all that is required.

This interpretation explains the observable effects of habituation in the preferential looking based experimental method. Tracking positional change is initially interesting and rewarding, but with repetition the procedural memory associated with visual scanning strengthens and reward and interest decline. This is also likely when featural information is bound to the representation so that the specific motor action can be repeated given future encounters with the same stimulus.

This interpretation is intended to clearly demonstrate the significance of having no object knowledge. By removing all of the biases that result from assumptions, otherwise self-evident questions become important. What does it mean when the ball is occluded by the box? Why is a ball on a continuous trajectory necessarily perceived as an individual object? There is no innate reason why any two instances of the ball in different positions should be perceived as a single entity, and there is no reason why a ball disappearing from existence would be unexpected to the perceiver. If the system has no foundational object knowledge, a trajectory consists of a series of unconnected spatial changes over time, each of which could conceivably be perceived as a new instance of the surface. Perhaps, rather than a ball in motion along a single trajectory, to an infant the scene is perceived as many instances of a ball on different trajectories, each appearing and then disappearing from existence.

### **Contribution of Interpretation 2**

The interpretation of experiment 2 introduces the issue of multiple interesting surfaces within a scene, and the learning required for the cohesion of similar but disparate positional change. Based on the eye tracking data, an argument is made that younger infants receive incomplete sensory input if there is too much positional change occurring at any one time. Infants initially perceive different regions of the scene separately and must learn to view it more holistically. This argument is based on the idea of episodes being short and specific events that gradually accumulate into larger ones with repetition, as was discussed in interpretation 1 and 2.

The effect of visual scanning on tracking multiple positional change is also identified. If a surface is being followed along a predictable trajectory it becomes less rewarding to track because the amount of positional change associated with the surface decreases.

Essentially the target surface being tracked is similar in positional change to a stationary scene. This allows further interesting regions of the scene to be investigated by switching back and forth between the positions predicted by the sensorimotor associations.

### **Contribution of Interpretation 3**

The ability to track multiple spatial positions on a surface is used to support the interpretation of experiment 3. A spatiotemporal completion task similar to that of experiment 1 is used to demonstrate how surface segregation could operate and be learned in two-dimensional space. The interpretation uses the distinction between the episodes involved in occlusion and disappearance events to learn surface segregation. It is argued that young infants do not perceive an occlusion event in enough temporal detail to notice a difference between occlusion and disappearance.

The difference between the two is a partial occlusion event in which part of the surface is no longer visible. This change to featural information causes an inconsistency in the representation and a comparison against the previous episode that was on the same spatiotemporal trajectory. The partial occlusion event causes multiple regions of the surface to become important; in the case of experiment 3, the leading and the trailing edge of the ball.

This is the first example of learned object knowledge in the model that must consist of recorded featural information. The ball, half ball, and box in experiment 3 are all important. It is learned that the box is part of the occlusion event which is why this perceptual mechanism is labelled as surface segregation rather than occlusion – the two surfaces are individuals that interact; this doesn't imply any understanding of in front or behind.

### **Contribution of Interpretation 4**

Interpretation 4 demonstrates the issues associated with individuating and identifying multiple occluded objects within the representation. The evidence from an object individuation task suggests that limited information is available for recall from memory. Infants struggled to distinguish between two individual surfaces on different trajectories.

The naïve model therefore assumes that spatiotemporal and featural information is only available for the last seen surface. Two separate trajectories that are seen one at a time may only be perceived as a single trajectory even if the featural information differs between the two. This doesn't add any new perceptual mechanism to the model but it does support the claim that sensorimotor associations are the primary form of representation, and suggests that perception could be based on partial models of the world.

#### **4.6.2 Perceptual Primitives**

The objective of the naïve model is to explain experimental evidence using perceptual primitives. These haven't been explicitly described during the naïve interpretations because a strict definition of them is complicated and involves a collection of components. The perceptual primitives are listed and justified as novel using comparisons to their conceptual level equivalents.

##### **Continuity**

Spelke's (1994) definition of object continuity will be recalled from chapter 2 – "objects move on connected, unobstructed paths". So long as an object is following a trajectory it remains the same instance over time. An example of this assumption is Meltzoff and Moore's (1998) representational persistence model which requires object continuity to keep track of the spatiotemporal information related to an object's trajectory.

The naïve model provides a perceptual explanation for continuity. The spatiotemporal information that is bound to an object is the result of a procedural memory of the motor actions required to visually scan the stimulus as it moves. In the naïve model, a surface remains the same instance only while it is in focus. This is defined as an episode. Initially, an infant is only capable of recording short episodes because tracking a moving surface is a learned skill. With training, this skill becomes an automatic response.

Interestingly, this doesn't require mental representation. The motor actions associated with visual scanning are initially not bound to the featural information of a surface. Instead, the detection of positional change is all that is required to build up a procedural memory. This procedural memory can be applied to any positional change in the scene.

With repetition featural information is bound, creating a more specific sensorimotor association which can be applied to the same surface in the future.

In the naïve model, continuity does not suggest that the infant is able to maintain a representation of a single object so long as it remains on an unobstructed path. Instead, continuity is the learned skill that allows an infant to follow the same surface as it changes position over time. There is no object understanding, only a reflexive response to a stimuli.

### **Cohesion**

Spelke's (1994) definition of object cohesion will be recalled from chapter 2 – “objects move as connected, bounded units”. Fields (2013) argued that the categorisation of sensory information as being a “physical object” takes a longer period of time than the cohesion process. Therefore, general knowledge about objects is a post-hoc linguistic understanding at best, and has no bearing on the mental representation of objects.

The naïve model offers a perceptual explanation for cohesion. It follows from the explanation of object continuity. Continuity is the tracking of a single region of a surface changing position over time. Cohesion involves the tracking of multiple regions of an individual surface or a collection of surfaces that share similar positional change. As with continuity, this doesn't necessarily require featural information to be bound.

This process doesn't require that the regions of a surface are changing position in exactly the same way. In fact, the description of the perceptual mechanism suggests that it may be easier to achieve cohesion if the movement amongst surfaces is varied. An example of this is point-light walkers which consist of many different patterns of motion; at least one for each leg and arm plus one for the head and another for the torso. These all follow a predictable pattern of motion. When one pattern is learned, attention can be directed towards another because the novelty of that motion is greater than the novelty of the learned pattern. If the new pattern is different from the learned pattern, then it is more likely that this will be novel and interesting.

In the naïve model, cohesion does not suggest an understanding that multiple patterns of motion are part of the same object. Instead, like continuity, it is a reflexive procedural memory that is a best effort prediction of disparate patterns of motion. As with continuity,

if the pattern of motion cannot be predicted (random motion) then the reward associated with following it versus other random motion never changes and cohesion doesn't occur.

### **Contact**

Spelke's (1994) definition of object contact will be recalled from chapter 2 – “objects affect one another's motion if and only if they touch”, and Mandler's (2012) conceptual primitive contact which “refers to one object touching another”. Object contact is at a higher level than continuity and cohesion, and consequently the current iteration of the naïve model is unable to explain it using perceptual primitives alone. However, some of the foundation for this primitive can be explained. An understanding of contact requires that two surfaces are identified as individuals, and that those two individuals cannot simultaneously occupy the same position in space.

The surface segregation mechanism is arguably the first step towards the understanding that two surfaces are individuals. Prior to learning the series of episodes that results in occlusion, the disappearance of a surface is novel but essentially meaningless. The surface segregation mechanism results in the understanding that the interaction of two surfaces causes a qualitative change to the featural information associated with one of them. This doesn't imply an understanding of in front and behind.

It is argued that the naïve model provides a foundation for the contact primitive because after acquiring the segregation mechanism all that is required is to learn what happens when the trajectories of two surfaces intersect and one is not occluded by the other. If an infant were to experience surface contact prior to segregation, all that could be learned is solidity – the understanding that two objects cannot simultaneously occupy the same position in space. However, with knowledge of surface occlusion, the infant can learn that two seemingly similar trajectories, one contacting and another occluded by a different surface, are qualitatively different. This is prerequisite knowledge for depth from motion parallax, and the equivalent of Mandler's (2012) MOVE (BEHIND) conceptual primitive.

### **4.6.3 Bottom-Up Object Knowledge**

The explanations for continuity, cohesion, and contact do not constitute learned object knowledge. Instead, like all learning during Piaget's sensorimotor stage of infancy, they

are learned responses to perceptual stimuli. Object knowledge is typically defined as a set of abstract rules that can be applied in a variety of situations. It could be argued that the detection of positional change is a generalised mechanism; however, in this model it is not learned, and it remains nothing more than a reflexive action that on its own provides no additional significance to the perceiver. In contrast to conceptual object knowledge, this model has described early perceptual primitives that do not require any understanding.

In this way, the perceptual explanations for continuity, cohesion, and contact could be considered implementations of Spelke's (1994) object principles without any conceptual understanding. These implementations are then supported by Fields' (2013) critique of object persistence, as discussed in chapter 2, in that they do not require the innate understanding of the conceptual category "physical object". Instead, an object is self-describing; the perception of an individual surface moving along a trajectory doesn't require that the corresponding surface adhere to any pre-existing abstract rules. Essentially, the post-hoc linguistic definition of "object" that an adult understands is unnecessary.

However, developmental theory is about the development of knowledge, so it follows that for this model to be of more use it should demonstrate some developmental progression. How do the explanations for continuity, cohesion, and segregation develop into conceptual object knowledge? The novelty of the bottom-up approach is that development is theorised without conceptual understanding. Unfortunately, the reason for the novelty of this approach is that it is very difficult to explain development without using generalised conceptual knowledge as a foundation. A good way to visualise this lack of generalisation is to think about the perceptual abilities of a classical computer. Sensory input is a series of meaningless pixels undergoing different patterns of change. Perceptual development can only be achieved through learning patterns of change over time.

An example of bottom-up development is the segregation mechanism. An occlusion event is learned as consisting of three episodes occurring on the same trajectory; zero, partial, and full occlusion. This only requires a very short-term memory to compare changes to featural information over time. Of course, this is not generalisable knowledge. It is specific to the current situation. Further, its only purpose is to differentiate occlusion from

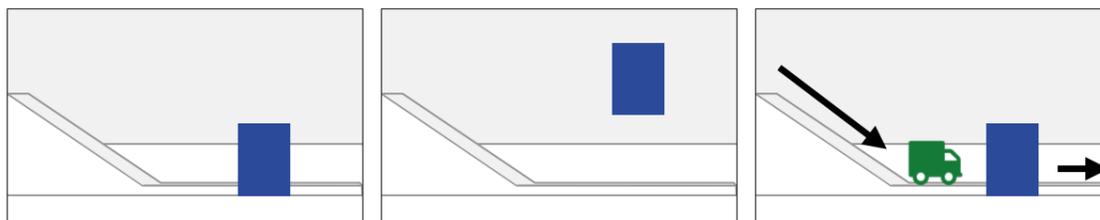
disappearance. There is not any meaning attached to the difference between the two; occlusion is simply an event that takes more effort to process and so is more interesting. This theoretically creates the opportunity for meaning to be attached later, for example once an understanding of in front and behind emerges.

This is where this model differs in function from the conceptual models of Mandler and others. The perceptual events that occur during occlusion are analysed and theorised about, but a description or explanation about any conceptual understanding to do with occlusion is avoided. As an example, recalling the interpretation of experiment 1, the infant is unaware that anything exists behind the box because this requires the conceptual understanding of at least occlusion and perhaps permanence. Instead, the naïve interpretation is that the infant learns the pattern of perceptual events in order to make anticipatory eye movements towards the opposite side of the box from where the ball was occluded. A conceptual understanding of occlusion and permanence can be added to this pattern of events at some point in the future once it is acquired.

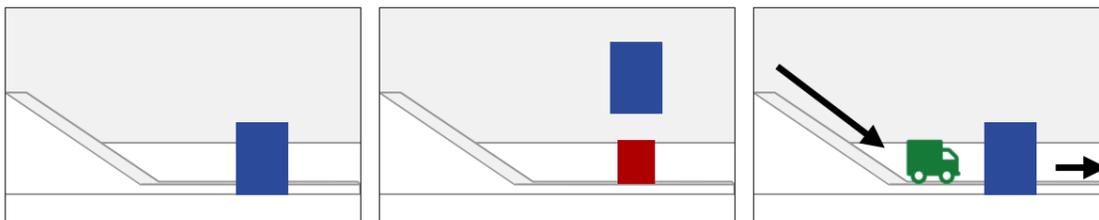
#### 4.6.4 Developing the Model Further

A fifth experiment is discussed to illustrate the primitives that are missing from a three-dimensional representation using the current iteration of the model. It also demonstrates that the current set of perceptual primitives are applicable in situations that differ from the experiment in which they were originally described.

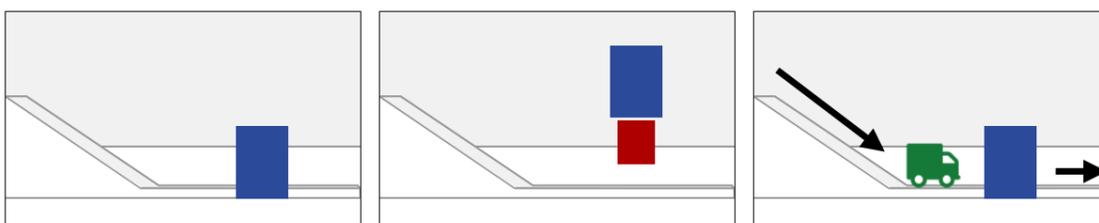
##### 4.6.4.1 Experiment 5 – Object Permanence



*Figure 10: Object Permanence Habituation Events. Left: Ramp and occluder. Centre: Occluder raised to reveal clear track. Right: Car rolls down ramp behind occluder. Adapted from Baillargeon (1986).*



*Figure 11: Object Permanence Impossible Test Event. Left: Ramp and occluder. Centre: Occluder raised to reveal blocked track. Right: Car rolls down ramp behind occluder. Adapted from Baillargeon (1986).*



*Figure 12: Object Permanence Possible Test Event. Left: Ramp and occluder. Centre: Occluder raised to reveal clear track with block behind. Right: Car rolls down ramp behind occluder. Adapted from Baillargeon (1986).*

This experiment is designed to determine whether or not an infant perceives a stationary object as continuing to exist behind an occluder. Infants are habituated to the display depicted in figure 10. The trial begins with a stationary ramp and occluder. The occluder is raised to reveal a clear path behind, and is then lowered to once again occlude the track. A car is released from the top of the ramp, rolls down the track on a trajectory behind the occluder, and continues to the opposite side of the display. After habituation the infant is shown four successive pairs of test displays. The test displays are identical to the habituation display except that a block is inserted behind the occluder. One test display is a possible event where the block is placed behind the path, as depicted in figure 12. The other test display is an impossible event in which the block is placed on the track, as depicted in figure 11. The interpretation of results is that if the infant looks longer at the impossible event, then the car and block are perceived as colliding behind the occluder, and therefore, the re-emergence of the car on the opposite side of the occluder is a violation of expectation.

#### 4.6.4.2 *Results and Original Interpretation*

Six and eight-month-old infants were tested. Both age groups looked longer at the impossible than the possible event. This result was interpreted as demonstrating that infants of these ages understand that an occluded object continues to exist behind the occluder. Additionally, infants can apply this knowledge to occluded trajectories and stationary objects. Infants are also able to apply knowledge of object solidity and contact to the representation and simulate the collision of a moving and a stationary object. This implies that the representation is encoded with featural as well as spatiotemporal information.

The first experiment was followed by a second in which the possible event consisted of the block in front rather than behind the path of the car. The rationale being that this would determine whether or not the display was being perceived literally and the impossible event was only more interesting because the block was closer to the infant. Again, both age groups looked longer at the impossible than the possible event. This study was repeated on four-month-old infants but the results were arguably less conclusive with females looking longer at the impossible event and males looking longer at the possible event (Baillargeon & DeVos, 1991).

Age	Interpretation of Result
6 & 8 Months	Infants look preferentially at the impossible event. Therefore, the infant perceives the block as existing behind the occluder and expects the car to contact the hidden block on its trajectory behind the occluder.

*Table 7: Summary of original interpretations of results for experiment 5*

#### 4.6.4.3 *Discussion*

The occluding screen slides up and back down revealing a clear track behind. The car then rolls down the ramp, behind the occluding screen, and to the other side of the display. With repetition, a sensorimotor association is created so that the trajectory of the car is automatically followed by the perceiver even across temporary occlusion. Habituation occurs over time as this sensorimotor association, and another created to follow the

occluding screen as it moves are learned. The positional change within the display becomes expected.

During the test trials the occluding screen is raised to reveal a block either on or behind the track. Otherwise, the positional change within the scene is equivalent to that of the habituation trials. Infants seem to notice the change from habituation to test, meaning that the representation of the habituation trials must include more than the positional change within the scene. The background behind the occluding screen must also be maintained to some degree, possibly as the result of a surface segregation mechanism similar to that of interpretation 3.

During the test trials infants look preferentially at the impossible over the possible event suggesting that further knowledge can be applied to surfaces within the scene. This requires a more substantial representation of the scene than is demonstrated by four-month-old infants in interpretation 4. In that interpretation it was contended that infants could only maintain a complete representation of the surface that was last seen. In contrast, it seems that in this experiment infants can maintain a detailed representation of both the block and the car.

This experiment is fundamentally different from the previous four in that the display is three-dimensional. The previous experiments all used a computer monitor as the presentation tool, while the display in this study was built using physical objects in the real world. Indeed, this experimental procedure relies on the perception of three-dimensional space; an infant sitting in a single position in front of the display must distinguish between a block on the track, and a block behind, or in front of the track.

The basic elements of this experiment can still be explained using the foundation formed by the previous interpretations. The sensorimotor associations of the first interpretation are demonstrated in the habituation to the occluder and car. The occlusion-based segregation mechanism of the second and third interpretations is plausibly in operation during the occlusion events. It seems that these perceptual primitives are generic enough to be applicable to situations other than those described in the original interpretations.

However, a complete interpretation of experiment 5 requires more information than can be collected from the perceptual primitives described thus far. The ultimate objective of this research should be to create a three-dimensional model of the world from the two-dimensional information arriving as sensory input. Marr (1982) built a similar model that includes the representations necessary to go from luminance intensities as sensory input, to an orientation and depth map, and finally to a hierarchical structure of three-dimensional objects.

Marr's model is a computational theory and does not stress the importance of goal-directedness to perception. In contrast, Gibson (1986) argued that visual information could be directly picked up and used by the perceiver so long as a reason exists to do so. The "ecological" approach contends that vision is a tool that enables animals to accomplish goals such as finding food and avoiding predators – a claim that seems evolutionarily plausible. However, Gibson's proposals for the algorithmic descriptions of these direct perceptual mechanisms were often unclear (Glennester, 2002).

There are similarities between the debate of Marr and Gibson, and the empiricist and constructivist debate. Marr believed that vision is an information-processing task and so can be reduced and reverse engineered to a representational foundation. Gibson believed, as Piaget did with young infants, that perception is based on goal directed motor action.

The naïve model already borrows from both of these accounts of visual perception, and further development would continue to do the same. Currently, the naïve model uses patterns of two-dimensional positional change and a goal-oriented reward system, both of which co-operate to drive visual perception and action. Utilising both of these pieces of information seems important in order to generate a three-dimensional model that could explain the empirical data of experiment 5. On the one hand, it is important to identify and learn the patterns of positional change and their correlation with external events. The position of the block in relation to the viewer and to the track must be maintained in order to differentiate the impossible and possible events. On the other hand, it is also important to understand the association of perception and goal-directed action. In the case of this experiment, this refers to the interest in following the car as it moves behind the occluding screen and the preferential looking at the apparent inconsistency that occurs in the infant's

representation when the car does not collide with the occluded block. If patterns of positional change are all that is important, then this should not occur because the visible positional change (the movement of the car) is equivalent during both the possible and the impossible test trials.

One important perceptual primitive that is required for a three-dimensional model involves perceiving a surface as being in front of, or behind another surface. This requires an understanding that a smaller object positioned higher in the scene is likely further away than a larger object positioned lower in the scene. Currently the naïve model distinguishes between an occlusion and a disappearance event. There is not an implicit understanding that an occlusion event means that one surface is in front of, or behind another surface.

The concept of in front and behind does not necessarily require object permanence because as Meltzoff and Moore (1998) point out, a pointer to an object can exist while out of sight without necessarily meaning that the object continues to exist in the physical world. However, this doesn't explain the simulated collision that appears to occur in the perception of this experiment. At the very least it seems that the shape and size of the block would need to be maintained during occlusion. It is difficult to imagine how an infant would perceive a collision to occur should the spatial position of both objects be marked only by some abstract placeholder and not by any featural properties. An important factor of the perceived collision is that the trajectories of the surfaces intersect. If the surface was not maintained in the representation as a solid object, then there could be no point of intersection.

A further requirement for the explanation of experiment 5 is the ability to maintain a representation of an occluded surface. This can be assumed as working similarly to the representation of objects in interpretation 4. It will be recalled that the empirical evidence suggested a partial representation of the two different surfaces where the featural information is only maintained for the one that was last seen. An interpretation using the current iteration of the naïve model would need to explain the partial representation of the important regions of the scene. This nuance could actually prove beneficial to the development of object knowledge in the naïve model because the author currently

theorises that the partial representations have something to do with the creation of new representations; essentially, they allow the merging of similar information.

If development of the naïve model continues in this direction, then it is possible that experiment 5 could be explained without the conceptual requirement of object permanence – the understanding that an object continues to exist in the world after it is no longer visible. The model could conceivably be built using only a combination of patterns of positional change and goal-directed motor action, in which case the notion of object permanence is only added later as a linguistic explanation of events.

## 5 CONCLUSION

### 5.1 Summary

The initial motivation for this thesis was to find ways to improve computational models of development. It became clear that a fundamental issue is that these models inherit the innate object knowledge attributed to infants by developmental researchers. There generally are rational reasons for these decisions in both academic fields; however, this remains a research topic in need of further investigation. Chapter 1 provides an example of the questions about object knowledge that emerge when some of the unstated assumptions are thought about more deeply. This background is used to support the research objectives and question.

The research question limits the scope of the literature review of chapter 2. The literature review attempts to demonstrate that historically, developmental research has relied on interpretations based on innate object knowledge, and that a bottom-up approach to interpretation is required to investigate this more thoroughly. Related literature involving the re-interpretation of infant experiments are reviewed in order to demonstrate the lack of research that aims to theoretically explain the perceptual mechanisms behind early object knowledge.

The lack of related literature results in the exploratory methodological approach presented in chapter 3. It is reasoned that the research methods must remain relatively unconstrained to allow for new insight to emerge. To this end, the methodology consists of two phases. Firstly, an organised literature review and data analysis phase is used to discover relationships between object knowledge in infant experiments. Later, a portion of these are selected for re-interpretation using the bottom-up approach. The re-interpretation process itself remains unconstrained. To provide some consistency an iterative approach is used in which the entire research process can be restarted if new information questions the validity of current findings.

Chapter 4 presents the findings as a list of the most significant re-interpretations and the respective relationships between the theoretical perceptual mechanisms. A discussion of

the naïve model follows, including its significance when contrasted against similar conceptual models from the literature review.

The thesis concludes that the preliminary model developed using the bottom-up interpretive approach can offer insight into the unstated assumptions around innate object knowledge.

## **5.2 The Bottom-Up Interpretive Approach**

The first research objective of the thesis is to develop a bottom-up interpretive approach. It was decided that a new interpretive approach was necessary because of the fundamental differences between perceptual and conceptual interpretations of experimental evidence. The bottom-up approach was built using an iterative research process in which intermittent reflection helped to gradually improve the findings.

The foundation of the bottom-up approach borrows from Marr. Marr's (1982) computational level of analysis is used to frame the object-related issue, and his algorithmic level of analysis supports the development of a theoretical algorithm to resolve the issue. The algorithm is described by the input and output information, and the expected process of transformation to get from the former to the latter (Marr, 1982). The bottom-up approach expects that the input information consists only of a combination of positional changes within the scene and associated motor action; and that the output information is the empirically supported response to the experiment, typically in the form of preferential looking. A plausible process of transformation between the input and output information is then described.

One of the key findings of this iterative research process was that experiments must be interpreted literally. Useful insight is only possible if interpretation occurs at a very low level of perception; nothing can have abstract meaning. For this reason, the bottom-up approach puts a lot of emphasis on positional change. Changes in spatial position have no inherent meaning, yet they are interesting and so draw the attention of the perceiver. Patterns emerge by separating positional change into episodes of time and considering each as an individual and isolated event. These episodes can then accumulate into larger patterns in a sort of hierarchical structure. The patterns remain meaningless; however,

pattern repetition works to increase or decrease the attention directed towards certain regions of a scene; a function that seems fundamental to perception.

Concepts that cannot be explained using only literal sensory information should be avoided. The original objective was never to explain the conceptual understanding of object permanence. However, the researcher thought it may be possible to explain more low level concepts such as occlusion. It became clear that this wouldn't be the case. Instead, the focus of the bottom-up approach changed to uncovering more information from patterns of sensory input. In the example of occlusion, focus turned to describing the positional change within the scene that corresponds with an occlusion event.

Another important finding was that the empirical evidence from infant experiments is not always the most important aspect of an interpretation. Contradictory empirical data is widespread within the developmental literature. Often the cause is experimental procedures that are similar but slightly different to the original. The issue is compounded when trying to organise a timeline of development by age. For this reason, the re-interpretations should not be taken literally. They are logical inductions formed through the careful consideration of all of the variables that were available at the time. They are not meant to argue that infants at age  $x$  are not capable of conceptual thought; rather, they demonstrate that there is the possibility that the empirical evidence can be explained without this level of understanding.

Perhaps the most important feature of the bottom-up interpretive approach is that it attempts to emphasise and then avoid the bias associated with assuming that the infant is automatically capable of perceiving a scene exactly as an adult does. The model suggests that early problems of perception may in fact be problems of information collection, and that learning to observe the important regions of a scene is an essential foundation.

### **5.3 The Naïve Model**

The second objective of the thesis is to investigate early object knowledge at a perceptual level by using the bottom-up interpretive approach. Because of the exploratory nature of this research only one question was originally posed:

- Can a model built using perceptual primitives describe the empirical evidence from infant behavioural experiments related to the development of object permanence?

The general discussion of chapter 4 shows that further perceptual primitives are necessary to explain the empirical evidence of a difficult object permanence task. The interpretation of experiment 5 (section 4.6.4) is included to demonstrate the difficulty of the original research question, and to show that there is reason to continue the development of the naïve model.

As discussed in chapter 2, there are actual empirical studies that have repeated similar experiments with altered methods and found perceptual explanations for interpretations that originally attributed infants with conceptual knowledge. However, none of the reviewed literature was concerned with theorising the development of those perceptual explanations.

This is where the current research deviates from the majority of other developmental research. As discussed in chapter 2, the intention of this research is not to theorise a pathway between perception and cognition – one of the fundamental problems of developmental science. The contribution of this model is to provide an understanding of the patterns of perceptual events to which conceptual meaning could later be associated. Through the careful consideration of patterns of perceptual events, it is possible to determine the significance of various types of visual stimuli. It is this literal perception of a scene that is missing from nativist interpretation, and to some extent developmental research in general.

This is a somewhat controversial topic and the intention of the researcher is not to argue with or discredit the majority of developmental research. In fact, the naïve model should slide in and fill the gap left open at the foundation of conceptual models. A fundamental reconsideration of development would only be required if it turns out that the model can explain early conceptual knowledge; an outcome that isn't likely considering the limitations discussed at the end of chapter 4, most importantly that the learned sensorimotor associations are not generalised from one stimulus to another.

The research objective was to determine whether a bottom-up approach to perception provides more or less explanatory power in regards to the development of early object knowledge. The answer is complicated. The model provides more explanatory power in that there is little to none when object knowledge is assumed as innate; however, it is debateable whether the model actually explains the development of object knowledge. It certainly explains the progression of specific perceptual mechanisms, but these do not operate using the abstract and generalisable rules normally associated with knowledge. On the other hand, it could also be argued from a Piagetian standpoint that there is no early object knowledge, and instead only learned responses to perceptual stimuli. Perhaps, abstract object knowledge is merely a post-hoc linguistic explanation for an otherwise perceptually driven phenomenon.

Chapter 1 began with a discussion about intentionality, original thought, and the philosophical issue of how a thing comes to be about another thing. Topics that are often debated in the context of mental representation; if one thinks of an object, it exists differently in the mind than it does in reality. It is clear that the naïve model, at least in its current version, does not meet the requirements. Firstly, the naïve model does not allow the perceiver to consciously think about an object. Instead, memories of familiar scenes are automatically brought to mind in specific situations in order to support the operation of perceptual mechanisms. Further, the skill learning demonstrated is not about anything, and all mental representations are literal copies of sensory information.

Yet, it could be argued that the development of perceptual mechanisms does create information different from that which is input from the senses. Take for example, the three stages of occlusion as a surface moves on a trajectory behind another surface, as discussed in section 4.3. The literal view of this consists of the surface, half of the surface, and no surface successively on the same spatiotemporal trajectory. This information alone is meaningless; however, the processing that occurs as a result of the change to featural information leads the perceiver to look more closely at the position where the surface was last seen prior to disappearance. This in turn allows for the segregation of surfaces (see interpretation 3 of chapter 4 for details). It is unclear whether this provides an understanding of the individuation of the two surfaces, but it can be argued that it provides

qualitatively different information about literal sensory information. In this regard, the information associated with the occlusion event could be said to be about something other than the literal view of the event; this may not fit the classic definition of intentionality, but it could be considered foundational.

## **5.4 Novelty and Contribution**

This study developed a bottom-up approach to interpreting infant behavioural experiments. The approach can be considered novel because it differs to any of the related literature that was reviewed, and is positioned to provide insight to an often unstated issue in developmental research – the development of early object knowledge. This research was exploratory and the bottom-up approach resulted from several iterations of discovery and development.

A separate contribution was the model of development that resulted from the re-interpretation of several behavioural experiments. The author used a novel method of data collection, analysis, and interpretation to find relationships between some of the mechanisms responsible for visual perception.

The current iteration of the model is preliminary and opportunities for future improvement are numerous. However, the current model remains a potentially significant contribution to artificial intelligence and cognitive science as it offers a methodological approach to the computational analysis of infant behaviour.

## **5.5 Limitations**

This is exploratory interpretive research and as such there are limitations inherent to the approach. Firstly, the re-interpretations are the best explanations given the variables available at the time. Significant effort was made to improve the validity of these interpretations through an organised experimental selection process, and iterative development of the model and interpretive approach. However, it is plausible, maybe even likely, that some information important to a theory of early object knowledge was missed. Fortunately, issues like this do not automatically render this research unsuccessful. The point of exploratory research is not to attempt to prove what an infant is thinking in a

specific situation, but rather to generate new ideas for what the infant might possibly be perceiving.

The selection of developmental studies for re-interpretation was necessarily biased towards areas of research that included the most literature. Experiments were only selected for re-interpretation if they had interesting empirical data and potential relationships with concepts tested by other experiments. This was necessary in order to gather useful insight within the provided time limit. Even with this selection bias, the direction of the research changed several times. Research topics that were investigated but did not make it into the current iteration of the model included relative and absolute direction, size constancy, and optic flow. The addition of these and other concepts is important for the future development of the model.

## **5.6 Recommendations**

The objective of this thesis was to point out the potential issues associated with assuming innate object knowledge and to work towards a solution. The model requires more supporting evidence before it can be recommended as an alternative for current models of the development of early object knowledge. At this point, the only recommendation to researchers is to suggest thinking further about the implications of assuming innate object knowledge.

## **5.7 Future Work**

This thesis was concerned with developing a bottom-up approach to interpretation, and the foundations for a bottom-up model of object knowledge. Further development of the model was discussed in section 4.6.4. An important inclusion would be the perception of depth and three-dimensional space, which seems closely related with the concept of one surface being in front or behind another surface.

## 6 REFERENCES

- Allen, J. W., & Bickhard, M. H. (2013). Stepping off the pendulum: Why only an action-based approach can transcend the nativist–empiricist debate. *Cognitive Development*, 28(2), 96-133.
- Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in cognitive sciences*, 9(3), 92-98.
- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6-and 8-month-old infants. *Cognition*, 23(1), 21-41.
- Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. *Developmental psychology*, 23(5), 655.
- Baillargeon, R. (2000). Reply to Bogartz, Shinskey, and Schilling; Schilling; and Cashon and Cohen. *Infancy*, 1(4), 447-462.
- Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, 3(1), 2-13.
- Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, 62(6), 1227-1246.
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20(3), 191-208.
- Bardi, L., Regolin, L., & Simion, F. (2011). Biological motion preference in humans at birth: role of dynamic and configural properties. *Developmental science*, 14(2), 353-359.
- Bogartz, R. S., Shinskey, J. L., & Schilling, T. H. (2000). Object permanence in five-and-a-half-month-old infants? *Infancy*, 1(4), 403-428.
- Brentano, F. (1874). *Psychology from an Empirical Standpoint*.
- Byrne, A. (2006). *Intentionality*.
- Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy*, 1(4), 429-446.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*: Sage.
- Fields, C. A. (2013). The principle of persistence, Leibniz's law, and the computational task of object re-identification. *Human Development*, 56(3), 147-166.
- Franz, A., & Triesch, J. (2010). A unified computational model of the development of object unity, object permanence, and occluded object trajectory perception. *Infant Behavior and Development*, 33(4), 635-653.

- Gibson, J. J. (1966). *The senses considered as perceptual systems*.
- Gibson, J. J. (1986). *The ecological approach to visual perception*: Psychology Press.
- Glaser, B. G., & Strauss, A. L. (2009). *The discovery of grounded theory: Strategies for qualitative research*: Transaction Publishers.
- Glennerster, A. (2002). Computational theories of vision. *Current Biology*, 12(20), R682-R685.
- Guerin, F. (2011). Learning like a baby: a survey of artificial intelligence approaches. *The Knowledge Engineering Review*, 26(02), 209-236.
- Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, 21(2), 167-179.
- Hayes, P. J. (1985). *The second naïve physics manifesto*.
- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences*, 100(18), 10568-10573.
- Johnson, S. P., Bremner, J. G., Slater, A., Mason, U., Foster, K., & Cheshire, A. (2003). Infants' perception of object trajectories. *Child Development*, 74(1), 94-108.
- Johnson, S. P., & Náñez, J. (1995). Young infant's perception of object unity in two-dimensional displays. *Infant Behavior and Development*, 18(2), 133-143.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where Infants Look Determines How They See: Eye Movements and Object Perception Performance in 3 - Month - Olds. *Infancy*, 6(2), 185-201.
- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive psychology*, 15(4), 483-524.
- Kitchenham, B. (2004). *Procedures for performing systematic reviews*. Keele, UK, Keele University, 33(2004), 1-26.
- Krøjgaard, P. (2004). A review of object individuation in infancy. *British Journal of Developmental Psychology*, 22(2), 159-183.
- Kuhlmeier, V. A., Troje, N. F., & Lee, V. (2010). Young Infants Detect the Direction of Biological Motion in Point - Light Displays. *Infancy*, 15(1), 83-93.
- Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developing 'what' and 'where' systems. *Trends in cognitive sciences*, 2(1), 10-18.
- Lungarella, M., Metta, G., Pfeifer, R., & Sandini, G. (2003). Developmental robotics: a survey. *Connection Science*, 15(4), 151-190.

- Mandler, J. M. (1988). How to build a baby: On the development of an accessible representational system. *Cognitive Development*, 3(2), 113-136.
- Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological review*, 99(4), 587.
- Mandler, J. M. (2012). On the spatial foundations of the conceptual system and its enrichment. *Cognitive Science*, 36(3), 421-451.
- Mareschal, D., & Johnson, M. H. (2003). The “what” and “where” of object representations in infancy. *Cognition*, 88(3), 259-276.
- Mareschal, D., & Johnson, S. P. (2002). Learning to perceive object unity: A connectionist account. *Developmental science*, 5(2), 151-172.
- Mareschal, D., & Thomas, M. S. (2007). Computational modeling in developmental psychology. *Evolutionary Computation, IEEE Transactions on*, 11(2), 137-150.
- Marr, D. (1982). *Vision: A computational approach*: Freeman & Co., San Francisco.
- Meltzoff, A. N., & Moore, M. K. (1998). Object representation, identity, and the paradox of early permanence: Steps toward a new framework. *Infant Behavior and Development*, 21(2), 201-235.
- Meltzoff, A. N., & Moore, M. K. (2001). Discovery procedures’ for people and things—the role of representation and identity. *Emerging cognitive abilities in early infancy*, 213-230.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*: Sage.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action (Vol. 27)*: England.
- Müller, U., & Overton, W. F. (1998). How to grow a baby: A reevaluation of image-schema and Piagetian action approaches to representation. *Human Development*, 41(2), 71-111.
- Newcombe, N. S., & Learmonth, A. (1999). Change and continuity in early spatial development: Claiming the “radical middle”. *Infant Behavior and Development*, 22(4), 457-474.
- Piaget, J., Cook, M., & Norton, W. (1952). *The origins of intelligence in children (Vol. 8)*: International Universities Press New York.
- Popper, K. (2005). *The logic of scientific discovery*: Routledge.
- Regal, D. M. (1981). Development of critical flicker frequency in human infants. *Vision Research*, 21(4), 549-555.

- Reiter, B. (2013). *The Epistemology and Methodology of Exploratory Social Science Research: Crossing Popper with Marcuse*.
- Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? *Developmental psychology*, 35(2), 427.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological review*, 107(2), 358.
- Schlesinger, M., & McMurray, B. (2012). The past, present, and future of computational models of cognitive development. *Cognitive Development*, 27(4), 326-348.
- Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, 22(5), 563-591.
- Siewert, C. (2002). *Consciousness and Intentionality*. Retrieved April 09, 2015, from <http://plato.stanford.edu/entries/consciousness-intentionality/>
- Slater, A., Johnson, S. P., Brown, E., & Badenoch, M. (1996). Newborn infant's perception of partly occluded objects. *Infant Behavior and Development*, 19(1), 145-148.
- Slater, A., Mattock, A., & Brown, E. (1990). Size constancy at birth: Newborn infants' responses to retinal and real size. *Journal of Experimental Child Psychology*, 49(2), 314-322.
- Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior and Development*, 13(1), 33-49.
- Slater, A., Morison, V., Town, C., & Rose, D. (1985). Movement perception and identity constancy in the new - born baby. *British Journal of Developmental Psychology*, 3(3), 211-220.
- Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, 50(1), 431-445.
- Spelke, E. S. (1985). Preferential-looking methods as tools for the study of cognition in infancy.
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, 10(1), 89-96.
- Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: Why we should no longer abide the nativist-empiricist debate. *Child Development Perspectives*, 3(2), 79-87.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27(2), 237-246.

- Valenza, E., & Bulf, H. (2011). Early development of object unity: evidence for perceptual completion in newborns. *Developmental science*, 14(4), 799-808.
- Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in newborn human infants. *Child Development*, 77(6), 1810-1821.
- Volkman, F. C., & Dobson, M. V. (1976). Infant responses of ocular fixation to moving visual stimuli. *Journal of Experimental Child Psychology*, 22(1), 86-99.
- Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, 72(2), 125-166.
- Wilcox, T., & Baillargeon, R. (1998). Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive psychology*, 37(2), 97-155.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358(6389), 749-750.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive psychology*, 30(2), 111-153.

## 7 GLOSSARY

<b>Surface</b>	The literal two-dimensional view of the world containing only the information available from the light arriving at the retina.
<b>Mental Representation</b>	A history of past information to which sensory input is compared. Consists of memories, ideas, and knowledge.
<b>Encoding</b>	The process of adding information to the representation.
<b>Object Knowledge</b>	The set of abstract rules that can be applied to objects.
<b>Perceptual Mechanism</b>	Object knowledge that performs an operation in the representation.
<b>Perceptual Primitive</b>	Object knowledge that can be described with perceptual information.
<b>Occlusion</b>	The effect of one surface blocking another surface from view.
<b>Disocclusion</b>	The opposite of occlusion. The effect of one surface revealing another that was previously blocked from view.
<b>Occluder</b>	The surface that is occluding another surface. In developmental experiments the occluder is typically a screen.
<b>Object Permanence</b>	The conceptual understanding that a hidden object continues to exist in the world after it is out of sight.
<b>Sensorimotor Association</b>	A form of representation in which sensory information is bound to motor action.