
by

Peter W. Coppin

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Faculty of Information
University of Toronto

© Copyright by Peter W. Coppin 2014
Abstract

Right now you are reading a sentence. Earlier, you might have been looking at a realistic picture, such as a photograph, or an outline drawing in a set of instructions. If you are a programmer, you work with sentence-like structures, such as code, or a system diagram. These are all graphic representations. To varying degrees, the effectiveness of every graphic representation relies on its ability to convey the designer’s intended meaning and elicit the intended reaction from its audience.

However, the design of graphic representations, even in technical domains such as visual programming language design or interactive information visualization, currently relies heavily on general principles based solely on practice, intuition, and informal measures of effectiveness from the applied art and craft of design (as opposed to scientific analysis or theory). There is an increasing demand for a scientific understanding of design and its evaluation from stakeholders (who seek evidence for effectiveness) and designers (who seek to advance their field). Because both the creation of graphic displays and their perception are literally embodied experiences, a model was developed with an embodiment orientation, specifically based on how graphics are perceptually and cognitively processed.

In my research, I found that graphic representations are constituted of two properties, pictorial and symbolic information, that emerge through two interrelated aspects of perception. In sighted individuals, for example, every graphic representation makes use of biological capabilities to process visual sensation (i.e., light hitting the retina), which are processed in
relation to culturally-learned capabilities (i.e., writing). I observed how graphic representations – such as pictures, diagrams, and sentences – are “naturally selected” (i.e., during different phases of design or problem solving). From these observations, I developed a model that distinguishes and predicts the effectiveness of pictures, diagrams, and sentences, in terms of how object relations and attributes are pictorially or symbolically represented, relative to the functional roles of those representations, contexts, and in some cases, individual perceptual-cognitive differences among perceivers.

This model is a step toward a science of graphics that could lead to evaluation techniques for information systems, theories for inclusive design, and ergonomically designed software programming tools.
Acknowledgments

I would like to thank several groups of individuals who contributed to this project: doctoral co-supervisors, the rest of the dissertation committee, the Faculty of Information, the Knowledge Media Design Institute, other participating labs and groups at U of T (and Toronto more widely), those who were part of projects that inspired me to pursue the research objectives that brought me here (back at Carnegie Mellon University and in Pittsburgh more widely), mentors, friends, and colleagues who influenced my world view back in my Texas and University of Dallas days, the friends and family who were with me the entire time, and most of all my, immediate family members who facilitated such an educationally rich environment when I was growing up. This section is the last item I am writing prior to submitting my dissertation and so it is likely that I might accidently leave someone out! Thus, the ‘final’ version of this section will be posted at www.petercoppin.org/academic/acknowledgments.

Co-supervisors. First, I would like to thank my doctoral co-supervisors: Dr. Stephen Hockema and Dr. Brian Cantwell Smith. My conversations with Steve began during my first days in the program and became a resource that aided my navigation of ideas from cognitive science, information theory, philosophy, linguistics, computer science, and beyond. I also want to thank Steve for his feedback on many drafts, and for his ability to appreciate how the idiosyncratic approach I brought to the project could ultimately produce results. I am glad that we are actively formulating ways to continue these conversations into the future. Brian was also a key part of these conversations, challenging me to communicate the dissertation’s concepts more clearly and responding with a rare kind of feedback that revealed his uncanny capability to interact with theories and text. Brian also helped me identify that an early outline I had for the model presented in this dissertation could be the central focus of this project. These conversations with Steve and Brian were the highlight of my experience as a PhD student.

Brian also deserves thanks for contributions that he might not be fully aware of: In many ways, he fostered a structure that allowed aspects of the project to hold together. He was Dean of the Faculty of Information when I interviewed at the iSchool and during the first years of my PhD. In this role, he cast a vision that drew many individuals to the research environment where my research transpired. Even when I was not directly involved in their conversations and activities, themes from their interactions were part of my conversations with Steve and Brian.
I would also like to thank Dr. Lynne Howarth for joining the committee in 2010. Our interactions began in the research methods course she taught during my first year in the program. She facilitated many rewarding conversations (and debates) with colleagues in my cohort. Many thanks to Lynne for helping me navigate the doctoral process (and for being such a practical problem solver). Lynne is a true role model and I hope to emulate her example as I move forward as an advisor and professor.

**External reviewers.** I would like to thank Dr. Corinne Jörgensen for expressing interest in my project after she saw my poster at the iConference in 2012 and for agreeing to serve as an external reviewer on the dissertation. Her feedback was detailed, insightful, and reflected the deep perspectives of a scholar who has considered issues that relate to my dissertation’s research. I would also like to thank Prof. Nick Wooldridge, a true practitioner of pictorial graphic representation, mastered through his many years of teaching and researching in the Biomedical Illustration program at U of T. Nick contributed thoughtful questions during the defense that inspired me to clarify parts of the dissertation. I would also like to thank Dr. Twyla Gibson for her participation and feedback on drafts prior to her departure to the University of Missouri.

**Cohort and other student colleagues.** I would like to thank fellow students from my cohort who were around to share ideas (and to help navigate the inevitable frustrations!): Lysanne, Lisa, and Lester. Special thanks to Lysanne for being a true collaborator by joining me (and assuming a proactive leadership role by) helping instigate the ‘Visual Thinking at KMDI’ lecture series and pro-seminar via KMDI (more about that below). Special thanks to Adam and Karen for the social interactions and conversations during my time at the iSchool. Special thanks for members of the C-Lab for our ideations and interactions.

**Knowledge Media Design Institute (KMDI) and Collaborative Program (CP).** Whereas the Faculty of Information was the ‘official’ home for this project at U of T, KMDI is the home of design at U of T, and was therefore also a home for my research. I want to thank Dr. Ronald Baecker, the original founder of KMDI who also served as the KMDI director during a formative phase of my research. Ron is another rare individual with a unique vision and an approach to digital media that reflects a true respect for design and the role that the social sciences can perform in it. In particular, I want to thank him for supporting the ‘The Visual Thinking at KMDI’ lecture series back in 2009. I would also like to thank Dr. Jim Slotta, the head of the
Knowledge Media Design Collaborative Program at the time for supporting our vision for the lecture series (and for participating as a guest lecturer in it). Dr. Barbara Soren also provided valuable administrative support, including the complex logistics that enabled guest speakers to join us Toronto. Special thanks to Prof. Nick Wooldrige joining us as the pro-seminar course that shadowed the lecture series and for working with Lysanne and myself to make it all happen. And of course, without Lysanne’s amazing efforts, I honestly do not think we could have convinced various decision makers and pulled it off. Finally, Special thanks to the guest lecturers: Neil Cohn, Sheelah Carpendale, Jim Slotta, Greg Van Alstyne, Nick Wooldridge, and Colin Ware. Their presentations and discussions created a ‘temporary research center’ that helped generate a conversation around graphic representation and cognition that heavily influenced my research.

U of T More Widely

Research groups. Many other groups at U of T also played an important role in this research project. I want to thank Dr. John Kennedy and his Perception Group at University of Toronto Scarborough for inviting me to join lab meetings and conversations. This was a first-hand introduction to ecological approaches to picture perception. Thanks also to members of John’s research group including Marta, Sheriff, Justin, and Marcello. I would like to thank Dr. Mark Chignell for inviting me to join his research group and brainstorming with me on several occasions when I was stuck. I also want to express my thanks and appreciation to both Dr. David Steinman and Dr. Dolores Steinman for our ongoing collaboration that continues to this day. Whereas my doctoral research developed a theoretical model, our collaboration is my first application of the model to solve a design problem.

Courses outside of KMDI and the iSchool. I would like to thank several professors who enabled me to either enroll in their courses (as a student from outside of their program) or unofficially attend lectures. These experiences helped introduce me to, and grapple with, ideas that I would not have encountered through other means. I was able to participate in the last few weeks of Dr. Kennedy’s picture perception course at UTSC and Dr. Morris Moscovitch’s course on higher cognition was timely: I had already sketched out the basic skeleton for the theoretical model that would become central to my dissertation over the preceding summer, but needed a broader introduction to perspectives from cognitive neuroscience that could be drawn upon to help support my thinking at the time. A term paper that I wrote in his course became the kernel
that I further developed for conference presentations and became the blue print for ideas in Chapter 2 of this dissertation. Special thanks to Dr. Karan Singh for letting me audit his computer science course focused on drawing interfaces. Thanks to Dr. John Vervaeke, I was able to attend many lectures in his introductory cognitive science course. This provided an additional perspective that helped me build a ‘map’ of theories that aided my journey to my own perspective. Thanks to Dr. Evan Thompson for the opportunity to attend some key lectures on the phenomenology of perception.

_Sponsors_. Several sponsors made it possible for me to focus on my doctoral studies. Special thanks for the National Science Foundation Graduate Research Fellowship Program (NSF GRFP) for funding this project for three years. I would also like to thank the University of Toronto and the Faculty of Information for fellowships that provided partial or full funding from 2007–2012 and for contributing travel funding.

_iSchool more widely_. I would like to thank Dr. Seamus Ross, who became the Dean of the Faculty of Information in 2009, for being so supportive, and for helping navigate the kinds of circumstances that inevitably arise over the course of a doctoral project. I would like to also thank administrative staff at the iSchool and the iSchool’s Infòrum for always being so helpful.

_Collaborations outside of U of T_. I would like to thank Dr. Jim Burton and his colleagues in the Visual Modeling Group at the University of Brighton for the informative conversations about diagrammatic reasoning and visual language systems. In particular, I would like to thank Jim for our fruitful collaborations. I am looking forward to more of these in the future.

_Past collaborations that inspired this project_. I would like to thank various collaborators from my Pittsburgh and Texas days. These early projects initiated the chain of thinking that led to this project. More specifically, I would like to thank everyone who worked on or was involved with EventScope, BigSignal, the Centre for Metahuman Exploration projects, and my early projects with colleagues from the University of Dallas. My experiences working on the ‘Life in the Atacama Project’ were particularly influential.

_Family_. I would like to thank Stephen and Sarah for the great conversations and for leading such interesting lives. Most of all, I would like to thank my parents for engendering an environment when I was growing up that valued education and creativity.
# Table of Contents

Acknowledgments .......................................................................................................................... iv

Table of Contents ......................................................................................................................... viii

List of Tables ............................................................................................................................... xiii

List of Figures .............................................................................................................................. xiv

List of Appendices ...................................................................................................................... xvii

1 Introduction ................................................................................................................................ 1
   1.1 Background, Problem, and Purpose .................................................................................... 1
   1.2 Context ................................................................................................................................ 4
   1.3 Proposition 35 of the *Elements*: A Site for Examining Pictures and Sentences .......... 8
   1.4 Foreshadowing the Result: A Model for Graphics ............................................................. 9
   1.5 Itemized Contributions ...................................................................................................... 10
   1.6 Significance ....................................................................................................................... 15
   1.7 Scope and Limitations ....................................................................................................... 17
      1.7.1 Scope of the Literature Drawn Upon .................................................................... 17
      1.7.2 Scope of Theoretical Explanation ......................................................................... 19
      1.7.3 Methodological Scope .......................................................................................... 20
   1.8 Outline ............................................................................................................................... 23

Part I: The Model ........................................................................................................................ 30

2 Pictorial vs. Symbolic Part A: Competing Accounts of Picture Perception from the Literature ......................................................................................................................... 31
   2.1 Preface ............................................................................................................................... 31
   2.2 Introduction ....................................................................................................................... 31
   2.3 Context ............................................................................................................................... 34
   2.4 An Applied Account: McCloud ........................................................................................ 34
   2.5 An Ecological Account of Picture Perception ................................................................. 36
2.6 A Constructivist Account of Picture Perception................................. 37
2.7 Discussing Ecological, Constructivist, and Applied Accounts............. 38
2.8 A Path to Reconciliation via the Dual-Route Hypothesis...................... 39
  2.8.1 Dual Route Hypothesis ................................................................. 39
2.9 Synergistic Aspects of Constructivist and Ecological Accounts........... 40
2.10 A Neuroscientific Account of ‘Recognition’...................................... 42
  2.10.1 Path to Reconciliation ................................................................. 43
2.11 Conclusions......................................................................................... 45
3 Pictorial vs. Symbolic Part B: Theory Development.............................. 47
  Preface........................................................................................................ 47
  3.1 Introduction: Perception-Reaction to Environmental Change and Variation 47
  3.2 Information (of a Graphic or Environment) as a Function of Perception 52
  3.3 Graphics as Intentional Visual Information........................................ 54
    3.3.1 Perceiving-Recognizing Random Environmental Changes Relative to an Organism’s Reaction: A Face Statue in a Forest ......................... 57
    3.3.2 Distinguishing an Intentional Reaction from a Reaction that is Intended to Cause an Intended Perception ...................................................... 58
    3.3.3 Pictorial Information ...................................................................... 62
    3.3.4 Symbolic Information ..................................................................... 64
  3.4 Conclusions......................................................................................... 65
4 Pictured vs. Symbolic Objects and Relations: Diagrams vs. Sentences .. 67
  4.1 Introduction......................................................................................... 67
  4.2 Prior Work ......................................................................................... 69
    4.2.1 Engelhardt: A Synthesis of Visual Information Design .................. 69
    4.2.2 Neural Basis for Graphic Objects and Spaces: Hurford ................ 70
  4.3 Pictured Objects............................................................................... 73
    4.3.1 Two Ways of Explaining .............................................................. 73
8.3 Broader Implications for Design Theory and Future Work........................................ 159
  8.3.1 Implications for Inclusive Design....................................................................... 161
  8.3.2 Broad Applications for the Model ...................................................................... 163
9 Epilogue: Dual-Coding Theory in Relation to the P3: Interference Prediction ............ 165
  9.1 Introduction.............................................................................................................. 165
  9.2 Dual-Coding Theory and Multimedia Learning: Sketch of Key Ideas...................... 166
  9.3 Comparing the MRP to Predictions 1–3 ................................................................. 166
  9.4 MRP Frustration in Computer Science Algorithm Education .................................. 167
  9.5 Conclusion: The Model is Synergistic with Dual Coding Theory and Can Also
Contribute to It ........................................................................................................ 169
10 References ................................................................................................................... 170
Appendices..................................................................................................................... 179
Appendix A: Proposition 35 of Euclid’s Elements ......................................................... 179
List of Tables

Table 1: An abbreviated version of the model developed in this dissertation. ................................. 10

Table 2: Deconstructing the example of over specification from Shimojima (1996)....................... 151
List of Figures

Figure 1-1: Graphics of Toronto’s Chinatown................................................................. 4

Figure 1-2: Itemized contributions are referenced to different parts of the model........... 11

Figure 1-3: The method (a) infers causes (b) for effects (c), whereas the causes for effects are interactions (d) between objects (e), such as a marked surface, and individual audience members (f) in order to afford an (effective) reaction (h) to a target situation (i).................................. 20

Figure 1-4: How the dissertation is organized............................................................... 25

Figure 2-1: Competing theories of this chapter presented on a spectrum..................... 32

Figure 2-2: A modified version of McCloud’s (1993) triangle........................................ 35

Figure 3-1: Pictures vs. symbols...................................................................................... 51

Figure 3-2: a) An environment, b) an environment that has been reacted to, c) an intentionally reacted to (configured) environment, d) an environment intentionally configured to cause an intended perception................................................................. 56

Figure 3-3: When Agent Sender sends a signal (1), Agent Sender can simulate (2) how Agent Receiver will perceive the signal (3). When Agent Receiver receives a signal (4), Agent Receiver can simulate (5) how Agent Sender simulated (6) how Agent Receiver would perceive the signal (7)........................................................................................................................... 60

Figure 3-4: Sharing perspectives through external representation................................. 61

Figure 4-1: An example of a diagram, which appears to have pictorial and symbolic properties.67

Figure 4-2: Graphic objects and spaces........................................................................... 70

Figure 4-3: A round green apple graphically represented via a picture, logic, and an English sentence......................................................................................................................... 73

Figure 4-4: Attributes as information that enable appropriate reaction to change-variation...... 81
Figure 4-5: Icon-buttons in Microsoft Word: Graphic Objects? .................................................. 82

Figure 4-6: Demonstrating pictorially represented relations among pictorially represented objects. .......................................................................................................................................... 87

Figure 4-7: Relations in information visualization. ................................................................................................................................. 87

Figure 4-8: A realistic picture as pictured relations among pictured objects................................. 90

Figure 4-9: A diagram as pictured relations among symbolic objects.............................................. 90

Figure 4-10: Symbolic relations among pictured objects. ............................................................. 91

Figure 4-11: Symbolic relations among symbolic objects............................................................... 91

Figure 5-1: More and less certainty. ............................................................................................. 96

Figure 5-2: Uncertainty and certainty applied to perception. ....................................................... 97

Figure 5-3: A conjunctive neuron. ................................................................................................ 99

Figure 5-4: The hierarchical organization of a perception-reaction system. ............................... 99

Figure 5-5: In this demonstration of the Concrete Pictorial Prediction (Prediction 2), a pictorially represented more concrete structure (a depiction of a more particular tree, left) symbolically represents a less certain, more abstract range of potential conceptual categories (right). .......... 104

Figure 5-6: The concrete structure of a coastline conveyed via pictorial properties of the outline drawings of a map (a). The concrete structure of an ocean floor conveyed via pictorial properties of the shaded regions of a map (b). ............................................................................................. 105

Figure 5-7: The art and craft of biomedical illustration specializes in the communication of concrete anatomical structures. This illustration is by Dorcas Hager Padget, who was a staff medical artist in the Medical Art Service Department at the University of Toronto in 1930..... 107

Figure 5-8: In this demonstration for the Abstract Symbolic Prediction (Prediction 2), the written word ‘tree’ (right) symbolically conveys a more certain conceptual category within which many different pictorially represented more concrete examples (left) could fall. ......................................... 108
Figure 5-9: The Interference Prediction (Prediction 3) anticipates that the cost of more pictorial information (and more concrete certainty) is less certainty regarding a conveyed abstract concept, which is manifested as less symbolic information. ............................................................. 114

Figure 5-10: An example of miscommunication caused by a graphic. .................................................. 115

Figure 5-11: Shared intentionality enables an author to anticipate how an audience might process pictorial and symbolic properties to receive intended concrete structures or abstract concepts. It reduces the range of possible audiences’ emulations that an author is required to simulate. ...... 117

Figure 5-12: Pictorial and symbolic information mapped to abstract and concrete concepts. ... 121

Figure 6-1: Babylonian clay tablet YBC 7289 with annotations. The diagonal displays an approximation of the square root of 2 in four sexagesimal figures, to approximately six decimal figures. Courtesy of Bill Casselman (http://www.math.ubc.ca/~cass/Euclid/ybc/ybc.html) and the Yale Babylonian Collection. ....................................................................................................... 127

Figure 6-2: Natural selection of sentences from a primordial soup of pictures. ............................... 137

Figure 7-1: Design problem-solving in interaction design. ................................................................. 147

Figure 7-2: Harry’s Memory Map from Shimojima (1996). .............................................................. 148

Figure 7-3: Another figure of Harry’s Memory Map from Shimojima (1996). ............................... 149

Figure 8-1: Sonifications of these roadmaps were produced by my students (Crosskey, Campbell-Smith & Kwok, 2013) ......................................................................................................................... 163

Figure 8-2: Potential research areas. .................................................................................................. 164

Figure Appendix A-1: Diagram used in proof of Proposition 35. .................................................. 179
List of Appendices

Appendix A: Proposition 35 of Euclid’s *Elements* .......................................................... 179
1 Introduction

1.1 Background, Problem, and Purpose

At this moment, you are reading a ‘sentence.’ Earlier, you might have been looking at a ‘realistic picture,’ such as a photograph, or the ‘outline drawings’ that sometimes accompany instruction manuals. If you are a programmer, you probably work with text-based interfaces to ‘computer code;’ you may also work with ‘diagrams’ of the system you are developing, perhaps using Unified Modeling Language (UML) or Entity Relationship Diagrams (ERDs). Herein, I use the general term ‘graphic representation’ or ‘graphic’ to cover all such representations and depictions (made in two dimensions).

This dissertation investigates how different types of graphics ‘afford’ actions. In design, an affordance is a relationship between an object and an individual that enables the individual to perform an action (Norman, 2002). For example, the shape and texture of a hammer handle affords a range of actions – driving a nail into a piece of wood, for example. In order to better anticipate and serve the needs and desires of their audience, good designers first understand – either explicitly or intuitively – how an artefact affords actions. For example, when designing the steering wheel of an automobile, a designer might anticipate how the physical properties of the steering wheel will interact with the anatomical properties of a driver’s hand and arm to afford a steering action.

How can the affordances of graphics best be explained and predicted in order to support visual interface design? Relative to the steering wheel example above, affordances of graphics fall somewhere between what might be described as perceptual (i.e., how the light of a graphic impinges on the retina) and physical (i.e., an action, such as when an individual grasps a mouse or clicks a button based on how a graphic dictates his or her action). Additionally, capabilities to perceive and act are shaped not only by physical, cultural, and linguistic differences, but also by experiences and events that transpire over a lifetime (James, 1890; Gibson, 1969). As a result, physical, cultural, and linguistic factors play a role in the affordances of graphics. A theory that explains the affordances of graphics would need to account for these factors.

Domains outside of the applied art and craft of design have traditionally investigated that which transpires between perception and action. For example, the ‘rational,’ ‘principled,’ or
‘scientific’ investigation of what transpires between perception and action has traditionally been the domain of philosophy, psychology, cognitive science, and neuroscience. Applied artists and designers, on the contrary, typically employ practical “rules of thumb” to guide design decisions in: cartography and graphic design (Bertin, 1983); visual information design (Tufte, 1983, 1990, 1997, 2006); and illustration (McCloud, 1993, 2006).

An example of current interface design practice that includes the applied art and design tradition can help expose the problem that this dissertation seeks to redress. Consider the contemporary information system design of: websites, computer games, eLearning applications, paper-based textbooks, blueprints, and instructions. In these systems, a user might see and interact with graphics such as photographs, instructional outline drawings, icons, labels, and sentences. The choice to include or design certain graphics for a given information system often depends on the particular skills of the development team members. Development teams often include experts in databases or computer programming, individuals with a humanities background who help create text or content for the site or system, and graphic representation designers who specialize in visual or graphic design and create user interfaces. The latter have often studied at an applied visual art or design school: designers learn how to draw graphic representations, arrange graphic elements of a display, and create computer programs to generate graphic displays. In design, less effort has been put into empirically or theoretically clarifying how and why people perceive, cognitively process, react to, and interact socially through graphics. Donald Norman, a cognitive scientist and pioneer in the field of human-computer interaction who is now a prominent design practitioner and theorist, recently highlighted this gap between applied art and design craft and the science of perception and cognition:

Classical industrial design is a form of applied art, requiring deep knowledge of forms and materials and skills in sketching, drawing, and rendering. The new areas are more like applied social and behavioural sciences and require understanding of human cognition and emotion, sensory and motor systems, and sufficient knowledge of the scientific method, statistics and experimental design so that designers can perform valid, legitimate tests of their ideas before
deploying them (Norman, 2010, Time to change design education section, para. 1).  

In other words, in the applied visual arts and design, little effort has been made to empirically assess the effectiveness of various graphic designs or to develop theories that predict how an audience will respond to design elements: do audiences respond to them in the way the designer intended?

Consider how a ‘science of graphics’ could aid interface design: Relying on rules of thumb may not be a problem when creating an effective painting or comic book. However, interface errors sometimes contribute to or even cause catastrophic events, such as nuclear disasters (Norman, 2002), plane crashes (Hutchins, 1995), spacecraft explosions (Tufte, 1990), and economic blunders (Soyer & Hogarth, 2012). In response to these errors, there is an increasing demand for theoretical models that incorporate scientific fields such as developmental psychology and cognitive science, theories of perception, and theories of graphic representation. For example, working in the computer science subfield of Information Visualization, Carpendale (2008) and identified the need for evaluation techniques that can be used to test and predict the effectiveness of information displays. Within the field of Information Systems, a domain that develops and uses ‘conceptual’ or ‘visual modeling’ languages, Moody (2009, 2005) noted that the majority of visual languages have not been evaluated to ascertain their usability and effectiveness, and that they were developed in the absence of a scientific theory or model that could inform design decisions. Ramadas (2009) identified similar problems within the field of educational psychology; she went a step further, aggregating previous research in developmental psychology, cognitive science, and the history of science to address issues of graphic representation, but has not yet developed a model that can be operationalized in design. There is a clear need for a scientifically grounded understanding of established conventions in applied visual arts to ensure that approaches to Visual Information Design can be better informed and more accountable.

This dissertation seizes the opportunity to redress the less systematized approach to the design of graphic displays graphic design with a more scientific one in order to clarify how

\[1\] Also see Ramadas (2009).
graphic displays afford actions. Specifically, it incorporates principles from fields of applied visual information design with those of perception and cognition to address the gaps in interdisciplinary knowledge. This kind of interdisciplinary approach will enable these seemingly disparate fields to illumine one another; it will also permit graphic designers to work with more information as they build graphic displays, and help them to evaluate the usability and effectiveness of displays. The purpose of this dissertation is to provide a perceptual-cognitive model that mediates or incorporates between theories from the fields of perception-cognition and design practices. More specifically, the project’s purpose is to develop a theoretical model that describes the perceptual-cognitive properties of the types of graphic representation frequently used in design (pictures, diagrams, and sentences), and to employ those properties to develop hypotheses regarding their affordances.

As a brief overview that locates this project’s research problem and purpose within relevant theories and design practices, let us imagine a short walk through Toronto’s Chinatown. In Chinatown, a rich ecosystem of graphics can be observed, exposing issues investigated in literatures across several fields of inquiry and practice including: art theory, analytic philosophy, perceptual psychology, human computer interaction, diagrammatic reasoning, and visual (programming) language design.

1.2 Context

When walking in this neighbourhood (Figure 1-1), I am surrounded by graphics that afford different actions. For example, pictures of food on various displays and menus help me
identify what food items I might purchase; however, because I do not read Chinese symbols, written graphics do not inform my actions. Nevertheless, the proprietors of these restaurants not only survive but indeed thrive in an urban environment where consumers from diverse linguistic and cultural backgrounds may or may not read Chinese symbols. Do the pictures on the menu boards afford actions across multiple linguistic backgrounds?

Goodman (1968), working at the intersection of analytic philosophy and visual art theory, claims that pictures are “linguistic” and that picture perception capabilities are “constructed” during the life of the individual (referred to as the “constructivist” account of picture perception). Perceptual psychologists from the ecological school of psychology disagree with advocates of the constructivist account, however, arguing that responses to pictures are not learned but arise from an individual’s biologically inherited or grounded capability to perceive and react within physical environments (referred to as the “ecological” account of picture perception; Gibson, 1968; Kennedy, 1974). The debate between ecological camps (in perceptual psychology) and constructivist camps (in visual art theory) remains unresolved (Kulvicki, 2010). In the visual arts, one of the dominant views is that a picture’s meanings are as arbitrary or constructed as the spoken words of a language. Yet, in the cognitive neurosciences, pictures are often employed as a proxy for the natural (non-represented) world in many experiments.

Despite these diverging opinions about how pictures inform individuals’ capacities and behaviours, the pressures of the consumer product market have led to design practices that are less theoretically based (relative to constructivist and ecological accounts of picture perception), but nevertheless reflect aspects of each competing view and therefore lend themselves to an integrated account that could inform design practice. In the Chinatown example, the menu and display designers must determine which aspects of their meals to represent with pictures, which to represent with text, how best to organize the display, which things to make prominent, what colors and other properties can do this, etc. They are operating within established conventions, influenced by their understanding of the cultural environment in which their displays will be processed, as well as, often, with a ‘target market’ in mind and a conceptualization of the needs and capabilities of such people.

‘Visual’ interface design. Professional designers and illustrators who work directly with audiences often employ design practices that reflect aspects of both constructivist and ecological
views. For example, McCloud (1993, p.13) suggests they operate on the assumption that pictures are “more easily received” insofar as audiences do not require specific learning or training prior to viewing and assimilating the content of pictures. Symbols, however, are presumed to be “less easily perceived” insofar as audiences do in fact require learning or training prior to being able to assimilate symbols. Indeed, McCloud’s view appears to be reflected in the design of many interfaces. Most computer users, for instance, are familiar with how GUIs are often described as ‘more natural’ or ‘more intuitive’ to use relative to the command line paradigm that the GUI surpassed in consumer markets. The ‘naturalness’ of a given GUI is often attributed to its visual or picture-like properties. Yet, a command line interface might also be considered visual insofar as there is light emitted by the computer screen’s graphics and picked up by retinal detectors of the eye. What, then, about a GUI is “more visual” than the non-visual command line of older GUIs? Furthermore, what is it about the visual interface that makes it seem more ‘natural’ or ‘intuitive’ to use for many audiences?

The practices of professional designers are reflected in the theoretical discourse surrounding interface design. For example, Norman (2002) urges designers to create a “natural mapping” between an interface and the inner workings of a system, claiming that the presence of written instructions is often a sign that a designer has failed to create such a natural mapping. Attempts to create a ‘natural interface,’ however, often through visual or picture-like graphics, has been far less successful for certain tasks, such as computer programming or the communication of logical concept in mathematics, where the potential limitations of picture-like graphics come to light and symbol-like graphics emerge as preferable.

**Where visual interfaces fail.** The personal computer might be one of the most reconfigurable devices ever to be mass-produced. A single computer can be programmed to perform a huge array of tasks that previously would have required a number of different devices to accomplish. Yet, only a small subset of computer users are equipped to create their own software programs; computers are difficult for most people to program (Pane, 2002; 2004). This has been characterized as a disparity between the design of the programming interface and the learning styles of computer users (Newell and Card, 1985; Pane, 1996).

The objective of trying to create programming languages that are easier to learn or use has become an active research area in the field of HCI (Kelleher & Pausch, 2003). One common
A strategy is to create so-called ‘visual programming languages’ (Boshernitsan & Downes, 2004). For example, the Alice project employs a 3D virtual environment to teach children how to program, and MIT’s Scratch interface is composed of visual icons that can be dragged into various ‘stages’ or other panels (Kelleher & Pausch, 2003; Boshernitsan & Downes, 2004). Unfortunately, visual programming languages have not made software programming measurably easier (Blackwell, 2006). For example, Fitter and Green (1979) studied flow charts and textual programming languages (Green et al., 1987) but found no evidence that either of these approaches increased their ease of use. Some evidence indicates that certain (graphic) representation systems are better for certain tasks, but none appear to be universally superior (Green et al., 1991; see Whitley, 1997).

Moody argued that the “principles for designing cognitively effective visual notations … ones that are optimized for human communication and problem solving” (2010, p. 485) are not adequately informed by scientific insights, such as how graphics are processed by a human perceiver. This is a key problem contributing to the usability challenges inherent to visual languages, and the lack of evidence to support any benefit from using these languages. Moody specifically focused on ‘conceptual modeling languages’ such as UML (Unified Modelling Language; Mishra, 1997) and i* (Intentional STrategic Actor Relationships modeling; Yu, 1995, 1997). Conceptual modeling languages are used for designing information systems (IS; Blackwell, 2006; Moody, 2009). These modeling languages are promoted as tools to ascertain the requirements of systems so that stakeholders, such as customers, operators, analysts, and designers, can better understand the system being modeled (cf. Borgida, Greenspan, and Mylopoulos, 1985). The usability of these modeling languages is poorly understood (Moody, 2006; 2009; 2010), and in the rare instances where it has been evaluated, it has been rated as poor (Moody, Heymans, and Matulevicius, 2009).2 Moody summarized the problem underlying visual language design as follows: “decisions about visual representation are typically made in a subjective way, without reference to a theory or empirical evidence, or justifications of any kind” (2010, p. 485).

---

2 Moody, Heymans, and Matulevicius (2009) was an evaluation of the i* (Yu, 1997) conceptual modeling language.
Mathematicians and logicians share the challenges faced by visual language designers in the field of diagrammatic reasoning. This field falls at the intersection of formal logic, visual programming language design, philosophy, perceptual psychology, and cognitive science. Though this field does not emerge from the applied art or design tradition, it contains a growing body of theories and empirical results that inform the theories developed in this project. In particular, Euclid’s controversial use of pictures in the *Elements* is a case study that I used to inform the theories developed in this dissertation, and I will use it within to demonstrate the explanations afforded by the model. The role performed by Euclid’s *Elements* in this project is introduced next.

1.3 Proposition 35 of the *Elements*: A Site for Examining Pictures and Sentences

In spite of the fact that geometry and its underlying theories are about spatial forms and relationships, geometric proofs are most often formally represented to people as text-based descriptions of the relevant geometric properties (Tennant, 1986). Not everyone agrees with the superiority of text-sentences in this context. Shimojima and Katagiri (2008) demonstrated that diagrams reduce ‘inferential load’ during reasoning by scaffolding visual-spatial aspects of memory. If this is true, why are proofs usually composed of text-sentences? In other words, why do mathematicians have less confidence in pictorial proofs? More specifically (and related to the theme artefact evolution which will be introduced in subsections below), if diagrams are so effective at reducing cognitive burdens during logical reasoning, then why were sentences ‘naturally selected’ for communicating logical structures?

Euclid’s *Elements* is a compelling site of artefact evolution for this project because its publication in 300 BC marked the birth of the axiomatic method that serves as the basis for most contemporary programming languages: understanding why proofs are ‘non-visual’ (from an ergonomic perspective) might help us understand why most contemporary programming languages are ‘non-visual.’ The axiomatic method that emerged was considered the pinnacle of argument and reasoning until it fell out of favour in the 19th century, precisely because of its reliance on pictures for key points in the argument (Mumma, 2008). The controversy about the method serves as a kind of ‘natural experiment’ as it involved debate specifically about text and pictures. Although the issue has not been debated within a cognitive science or visual art-design context, a close examination of the controversy can reveal other research that can form the basis
for an explanation. Because the axiomatic method and the text-sentences that seem to be associated with it (cf. Tennant, 1986) are the foundation of contemporary programming languages (Hoare, 1969; 1999), this explanation should be applicable to the human-computer interaction (HCI) problems introduced earlier in this chapter, which researchers are trying to address by developing more ‘intuitive’ (i.e., ‘less learned’ or ‘natural’) programming techniques.

1.4 Foreshadowing the Result: A Model for Graphics

To foreshadow the model that will be described in the chapters that follow, key aspects are now highlighted.

The theoretical model reported in this dissertation provides insight into why some so-called pictorial properties of graphics might seem to inform actions more ‘naturally’ e.g., with less effort, less learning, or less training. Additionally, the model provides insight into why some pictorial properties fail to afford certain kinds of actions (such as the communication of abstract concepts) that are more easily afforded by the so-called symbolic properties of graphics that might seem akin to spoken or written languages.

Pictorial properties. Simply put, what I will call the pictorial properties of graphics produce light structures that share common properties with light structures of a physical environment (cf. Gibson, 1968; Kennedy, 1974; Hammad, 2008). Individuals from diverse backgrounds develop abilities to perceive and act in physical environments: they develop ‘perceptual categories’ that are tightly coupled with sensory receptors such as the retina of the eye (Mandler, 2008). When an artist marks a surface, the reflected light is assimilated by a viewer who involuntarily defers to perceptual categories which have developed to enable the perception of surfaces and edges within a physical environment. As a result, the viewer is able to perceive pictorially represented surfaces and edges that are beyond, or other than, the marked surface.

Symbolic properties. The model also helps clarify why some concepts are difficult to communicate pictorially. Neural structures that engender what are commonly referred to as conceptual categories develop over a lifespan through the convergence of visual, auditory, and tactile sensory modes: they are more ‘amodal’ (Barsalou, 2009; Damasio, 1989). For example,
the concept of ‘sharp’ can be attributed to the indentation of a surface that is touched but not seen, as well as a drawn shape that is seen but not touched. In this situation, the symbolic properties of graphics exhibit many of the same learned and culturally specific properties of language. Concepts such as ‘hot’ and ‘spicy’ on a restaurant menu are afforded by the symbolic properties of written graphics. Concepts of logical reasoning and computer science are also afforded through symbolic properties of text-based programming languages.

**Hybrid graphics: Diagrams.** Some graphics, such as diagrams, are hybrids of pictorial and symbolic properties. My model distinguishes diagrams from pictures and text-sentences by describing how objects, relations among objects, and attributes of objects can be represented pictorially or symbolically.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Picture</td>
<td>Diagram</td>
<td>Text-Sentence</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Relation</td>
<td>Pictorial</td>
<td>Pictorial</td>
<td>Symbolic</td>
</tr>
<tr>
<td>C</td>
<td>Object</td>
<td>Pictorial</td>
<td>Symbolic</td>
<td>Symbolic</td>
</tr>
</tbody>
</table>

**Table 1: An abbreviated version of the model developed in this dissertation.**

### 1.5 Itemized Contributions

This section lists the specific contributions of this dissertation that will be discussed in the chapters that follow. Contributions will be referenced to this dissertation’s model in Figure 1-2.

The primary contribution of this dissertation is the model that distinguishes graphics (pictures, diagrams, and text) in terms of how each type of graphic affords the communication of
concrete or abstract concepts. More specifically, the model distinguishes and describes pictures, diagrams, and text in terms of how each type of graphic pictorially or symbolically represents objects, attributes, and relations.

Figure 1-2: Itemized contributions are referenced to different parts of the model.

Developing this model required developing additional supporting concepts. The core contributions could be thought of as the central claims of this dissertation, and the predictions that are made possible by these core claims. The core claims build on assertions that were synthesized from the literature on which this dissertation was built. Other supporting claims are also original to this dissertation, but are not core contributions. The supporting claims helped bridge the gap between work by other theorists and the core contributions.

A. Pictorial and symbolic information
Two core claims that are fundamental perceptual-cognitive properties of graphics are conceptualized in Chapter 3 as a means to distinguish different types of graphic representations (pictures, diagrams, and text): pictorial information and symbolic information.

The model distinguishes between pictorial and symbolic information based on two interrelated aspects of an individual’s homeostatic perception-reaction loop, a concept synthesized from the literature (see Chapter 2). I use ‘emulation’ to refer to the aspect of perception-reaction that ‘picks up’ or detects produced or reflected light that impinges upon sensory receptors (Chapter 3). I use ‘simulation’ to refer to the aspect of perception-reaction that processes, filters, or interprets that which is emulated, and thereby enables an organism to anticipate changes and variations that are relatively more distal, in terms of space and time (Chapter 3).

**C1: Pictorial Property/Information Claim.** Pictorial properties will be conceptualized as emulated properties of a graphic (Chapter 3).

**C2: Symbolic Property/Information Claim.** Symbolic properties will be conceptualized as pictorial properties that cause simulations of an author's intended meaning (structure). This ‘intended meaning’ can be used as a way to distinguish graphically represented items from non-graphically represented items (Chapter 3).

**B. Information of a graphic**

Conceptualizing pictorial and symbolic information also requires developing a way to conceptualize information in a graphic, in contrast to information in an environment. A (graphic) representation can be conceptualized as an environment that has been intentionally configured to produce (visual) information that causes an intended percept (Chapter 3).

**C. Predicted affordances of pictorial and symbolic information**

Sections A–B (above) enabled the following predictions:

**P1: Concrete Pictorial Prediction.** Pictorial information affords the communication of concrete concepts (more effectively than symbolic properties; Chapter 6).
**P2: Abstract Symbol Prediction.** Symbolic information affords the communication of abstract concepts (more effectively than pictorial properties; Chapter 6).

**P3: Interference Prediction.** Pictorial information interferes with the communication of abstract concepts via symbolic information (Chapter 6).

**D. A means to distinguish between pictures, diagrams, and text in terms of how objects or relations are pictorially or symbolically represented**

Building on the results summarized in sections A-C, three additional core claims of the model provide a way to distinguish between pictures, diagrams, and text in terms of how objects or relations are pictorially or symbolically represented. The model provides a way to conceptualize the following:

*C3: Composite Picture Claim.* Pictures are conceptualized as pictorially represented relations among pictorially represented objects (Chapter 4).

*C4: Diagrammatic Claim.* Diagrams are conceptualized as pictorially represented relations among symbolically represented objects (Chapter 4).

*C5: Sentential Claim.* Sentences are conceptualized as symbolically represented relations among symbolically represented objects (Chapter 4).

**E. Claims C3-C5 (above) require conceptualizing what it could mean for an object, relation, or attribute to be represented pictorially or symbolically (these are supporting claims):**

Objects can be conceptualized as targets for perception-reaction in relation to other potential targets for perception-reactions; attributes enable object-reaction targets to be distinguished (Chapter 4).

**F. Using the key terms summarized in section A-E (and the concepts to which they refer), it is possible to predict the following affordances of pictures, diagrams, and text:**

*P4: Composite Picture Prediction.* Pictures afford the communication of concrete relations among concrete objects (Chapter 5).
**P5: Diagrammatic Prediction.** Diagrams afford the communication of concrete relations among abstract objects (Chapter 5).

**P6: Sentential Prediction.** Sentences afford the communication of abstract relations among abstract objects (Chapter 5).

**G. Demonstrating contributions**

The strategy used to support/demonstrate the claims/contributions of this dissertation will be to show how predicted affordances correspond to functional purposes. I do this by identifying situations where we can be confident that the types of graphics observed have been naturally selected to effective states relative to functional purposes.

**Demonstration 1 focuses on P1: Concrete Pictorial, P2: Abstract Symbolic, and P3: Interference Predictions by examining the artefact evolution of Euclid’s Elements**

When the objective is to communicate an abstract concept, the model’s explanatory prediction is that symbolic properties will be more effective (will communicate abstract concepts with greater certainty) than pictorial properties (P1: Concrete Pictorial Prediction). Based on the principles of artefact evolution, symbolic properties are more likely than pictorial properties to be naturally selected when the functional purpose is to communicate abstract concepts. The P2: Abstract Symbolic Prediction is also demonstrated.

When the objective is to communicate a concrete concept, the model’s explanatory prediction is that pictorial properties will be more effective (will communicate concrete concepts with greater certainty) than symbolic properties (P1: Concrete Pictorial Prediction). Based on the principles of artefact evolution, pictorial properties are more likely than symbolic properties to be naturally selected.

These predictions are demonstrated by examining the history of Euclid’s Elements to explain how the text-sentences of the Elements were ‘naturally selected’ from a ‘primordial’ soup of graphical representations.

**Demonstration 2 focuses on the ‘Free Ride’ Phenomenon of Diagrams**
When the objective is to design a logical structure (such as a symbolically represented proof), the model’s explanatory prediction is that pictorial properties will be predominant in the early design phases, diagrammatic properties will be predominant in the intermediate phases, and symbolic properties will be predominant in the final phases. These predictions will be demonstrated by examining the “free ride” phenomenon often noted in diagrammatic reasoning communities (e.g., Shimojima, 1996).

1.6 Significance

Both fine and applied visual arts involve phenomena that would be missed by an account that ignores constructed ‘conventions’ such as traditional/accepted modes of presentation and interpretation through the history of art often created by artists themselves, shared understandings of political-religious/historical facts, tropes, common reference points, homage to inspirational landmark work, etc. However, less artistic uses of pictures may limit the purely constructivist account. For example, graphic artists may produce representations for educational materials; these representations must take into account individual similarities and differences (in terms of culture and perception – cognition). This kind of task may be explained in terms of cognitive neuroscience: students with dyslexia or autism, for instance, have a neurological configuration that differs from their ‘normal’ counterparts. Cognitive neuroscience research can help lead to more effective picture and symbol use in materials to accommodate differences; it can also help clarify which aspects of pictures can work across cultures. Therefore, a model that can include both ecological and constructivist accounts has more than academic implications: graphic representations are increasingly used in information displays, particularly in the fields of IT and education, so a principled scientific account that could inform the design of graphics is increasingly needed, as noted in previous subsections.

By clarifying what types and properties of graphic representation could afford the design of graphics, we can ensure design requirements for more useable representations (e.g., for logical system design). The next subsection will present one illustrative example of a potential benefit, in addition to enabling more useable languages for most computer users, by considering the complicating role of perceptual-cognitive diversity, or learning styles.

Contributions to accessibility. Consider the case of dyslexic computer science students. Powell (2004) reported that these students struggle with text-based programming environments,
but often excel in the highly-sought creative and lateral-thinking abilities required for Information Systems (IS) design. Powell’s observations are consistent with other findings regarding a non-standard set of text-processing weaknesses that correspond to visual-spatial strengths, for at least a certain subset of dyslexics (Schneps, 2007). Multi-modal, and/or ‘visual thinking’ techniques have been proposed as a way of scaffolding computer users with dyslexia to text-based materials (Schneps, 2007; Deibel, 2008) and textual-programming environments (Wilson, 2004; Coppin, 2008). Designing these representations, or developing strategies for using representations, appears to require: (1) a better understanding of how representations afford the design of logical systems for ‘normal controls,’ and (2) a better understanding of how these affordance vary across different learning styles. A representation system that enables a dyslexic to program would be an embodied response to the motivating question posed.

**Contributions to interface design.** Visual artists and designers are critical to many software design processes, and drawing and sketching plays an important role in the look, or feel, of a user experience (Buxton, 2007). This project could lead to design requirements for practices, representations, or software tools that facilitate the ecosystem between drawings and logical argument design. Such a system would also be yet another embodied response to the motivating question posed.

**Contributions to a science for visual information design.** A principled and scientific understanding of visual information design is needed, particularly in the domain of education (e.g., Mayer, 2003; Ramadas, 2009). This project could serve as a small step in this direction because each representation type that will be explored by this project is also widely used in education and beyond.

**Contributions to understanding the role of visual representation in problem-solving.** Other contributions are possible on a more philosophical level, and more generally. Simon wrote, “the proper study of mankind is the science of design” (1996, p. 159). This project is principally about design, but it also focuses on logical systems. Reasoning has been a topic of investigation since the dawn of philosophy. I do not seek to re-invent the wheel, but an analysis of the role that
drawings and other similar representations play in developing logical arguments may add some insights to these age-old questions.³

### 1.7 Scope and Limitations

For practical reasons, the scope of this investigation is limited in three ways: the scope of the literature drawn upon, the level of theoretical explanation, and the methodological scope.

#### 1.7.1 Scope of the Literature Drawn Upon

I limited the project’s scope by drawing only on the science and philosophy of perception to explain proto-theories of practitioners. While I was certainly influenced by several rich traditions outside of the science of perception, such as semiotics (e.g., Peirce, 1977; Saussure, 2011) and communication theory (e.g., Innis, 2007; McCluhan, 1994), I did not draw upon these traditions directly. I plan to reconcile the perceptual-cognitive theories developed in this dissertation with these other bodies of work in future research. In this dissertation, I limited the scope of the literature drawn upon to conceptualize pictorial and symbolic properties of graphics (Chapters 2–3) and to conceptualize how pictorial and symbolic properties can pictorially or symbolically represent objects and relations to produced diagrams and sentences (Chapters 4–5).

**Scope of Literatures for Pictorial versus Symbolic Properties**

Within the philosophy and science of perception, I further refined the focus of the project by synthesizing theories from a rare debate between the arts and sciences, specifically the debate between Goodman (1976) and Gibson (1986). Their direct ‘conversation’ exposed points of intersection and conflict, providing raw materials for the development of a model that is synergistic with the scientific ideas of ecological psychology and the more culturally-oriented perspectives from art theory, which were not addressed by ecological psychology.⁴ I then

---

³ Reasoning has been classically defined as a defining characteristic of human beings (Aristotle, 2001). Yet computers, a crowning achievement of rationality, are difficult to program for most computer users (Pane, 1996). Observing other, less rational, representations may contribute to revising this age-old issue.

⁴ My inspiration for this approach came from Bruno Latour’s presentation at the ‘Future of Objectivity Conference’ at the University of Toronto in 2008 (http://research.ischool.utoronto.ca/objectivity/).
compared my synthesis to the proto-theoretical explanations of working practitioners in illustration and graphic design (see e.g., McCloud, 1993), to produce an explanation that distinguishes between the pictorial and symbolic properties of graphics. To maintain this narrow focus, I did not draw upon the many rich theoretical traditions that lay outside this debate, such as Mayer’s (2002) multimedia theory of communication, or the mainstream cognitive psychology theories upon which Mayer’s theories are based, such as Baddeley’s (1975) theory of working memory.

By focusing on a debate about picture perception across art and science, I was able to expose traditions from experimental perceptual psychology, specifically ecological psychology, and to develop a biological account of perception that is synergistic with the affordance-based ideas of design (e.g., Norman, 2002). This biological account provides a way to explain and predict how individuals use perceptual capabilities inherited from their ancestors, which ground the learned capabilities that develop over a lifetime. However, much experimental psychology research focuses on perception of the ‘natural’ non-graphically represented world. By focusing on the debate between Gibson and Goodman, I exposed a small subset of experimental perceptual psychology that focuses on pictorial graphics (Gibson, 1978; Kennedy, 1974). However, Gibson and Kennedy were not trying to develop theories to inform design, or theories to explain the practices of designers in perceptual-cognitive terms. Thus, my objective was to extend theories from perceptual psychology to explain graphics, with a focus on ecological theories of perception that are synergistic with affordance-based ideas of design.

**Scope of Literatures for Diagrams versus Sentences**

Developing a perceptual account of graphics that could explain affordances of ‘language-like’ elements of graphics found in written graphics, such as text-sentences and the written labels of diagrams, entails the extension of ecological perception into new territory in Chapters 4–6. Graphics include many written ‘language-like’ elements that, though visually perceived, have not traditionally been the scope of (ecological) visual perception, because linguistics is treated as a separate field that draws upon different bodies of knowledge. Following the theme of artefact evolution, and an aim to develop theories that are compatible with design practice, I strategically selected a linguistic approach to graphics in Chapter 4 that was developed by Engelhardt (2002), who synthesized proto-theories of practitioners with linguistics and semiotics. I was able to
synthesize this linguistic account with the ecological model of Chapter 3 by drawing upon Hurford’s (2003) recruitment of Goodale’s (2005) dual-route hypothesis (introduced in Chapter 2). The result of the above was then compared with leading accounts of diagrammatic and sentential graphics form the diagrammatic reasoning community in Chapters 6-7.

1.7.2 Scope of Theoretical Explanation

The scope of the model development in this dissertation can be understood as one step toward developing a scientific model for a field that is traditionally practiced as a craft. By limiting this project to a dissertation scale, my goal was to produce a model that enables predictive explanations. I will demonstrate the model’s predictive explanations starting in Chapter 6, but experimentally testing the model is beyond the scope of this project and will be a focus of future work. Let us now consider what is required of the model.

One criterion for adequacy of the proposed model is that it must both emerge from, and accord with, phenomena observed in the field of graphic representation in such a way as to enable theories and explanations regarding their causes and effects.

A graphic representation should enable an individual who sees and understands it to react appropriately to, or act appropriately upon, the situation that the graphic representation was designed to enable the individual to react to. Therefore, the ‘effect’ engendered by a graphic representation can be defined in the following way:

• perceptual capabilities of individuals (including differences among individuals) allow
• properties of marked surfaces to be
• perceptually processed as graphic representations such that
• reactions are afforded to
• target situations in the environment.

Here, ‘target situations’ refer to what would classically be understood as the situations that graphic representations are considered to represent. More constructively, we can interpret these as the environmental situations to which the representations have been created to encourage responses in perceivers. Borrowing a term from Gibson (1986) and Norman (2002), I will call the reactions thereby engendered the ‘affordances’ of the representation. See Figure 1-1
for a schematic representation of the above, that relates the notion of inferring causal relationships (left panel) to affordances of graphic representations (right panel).

![Diagram]

**Figure 1-3:** The method (a) infers causes (b) for effects (c), whereas the causes for effects are interactions (d) between objects (e), such as a marked surface, and individual audience members (f) in order to afford an (effective) reaction (h) to a target situation (i).

**Relation to Type I, II, and III Levels of Theory**

Models have different levels of predictive power depending on the evolutionary phase of a field (Gregor, 2006). In Chapter 2, the synthesis of pre-scientific proto-theories from the applied arts, sparked by a famous debate across art and science, makes it possible to describe the properties of graphic representations at a Type I level (a theory or ‘language’ for analyzing, see Gregor, 2006). Further development, involving more recent findings from the cognitive neurosciences (Chapters 3–4) will help develop a Type II level theory (a ‘theory for explaining,’ see Gregor, 2006), and even venture into a Type III theory (a ‘theory for predicting,’ see Gregor, 2006, in Chapter 5). Finally, the predictive ability of the model will be demonstrated by explaining why and how conventions have been naturally selected to effective states (Chapters 6 and onward).

**1.7.3 Methodological Scope**

This dissertation adopts two primary methodological strategies. First, because of my commitment to derive the model from the lived practice of graphic designers, I will derive my observations (especially in the initial phases of this work) of the properties of graphic
representations from their use in naturalistic environments. Second, to address the lack of clarity endemic in such ‘in vivo’ situations, I will view these naturalistic settings in something of an evolutionary way, as ‘sites of artefact evolution’ (SAE) in which diverse representations types have evolved, and over time been naturally selected for their utility and worth.

**Sites of Artefact Evolution (SAE)**

*Characterizing SAE.* SAE can be defined as settings in which individuals configure artefacts of their environments to afford actions that contribute to prosperity and survival. Within this conceptualization, ‘artefact evolution’ is the process whereby an individual either replicates or modifies various types of artefacts (cf. Simon, 1985; Kirsh, 2009). Types or configurations of artefacts ‘survive’ if they are replicated, and types or configurations of artefacts that are not replicated ‘die.’

*Identifying ‘effective’ artefacts.* This subsection describes how SAE can be used to identify which types of artefacts may be effective, as a step toward explaining why and how types of artefacts may be effective. We can observe what types of artefacts survive, while also paying attention to the functional purposes that the artefacts appear to serve, as well as the overall context within which the type of artefact and functional purpose is located. If a type or configuration of artefacts survives, then we can try to develop a theory to explain why and how some of its properties might afford an action that contributes to its survival and prosperity (relative to a context and a functional purpose). Additionally, if a type or configuration dies, then the theory should explain why it did not afford actions that contributed to its survival and prosperity (relative to competing paradigms). For example, a car’s steering wheel is a type of artefact with properties that have remained consistent throughout the evolutionary history of automobiles. One functional purpose of a steering wheel is to enable a driver to point a vehicle in the intended direction. Driving to a destination is a context within which the steering wheel (a type of artefact) and its functional purpose (to direct the car) are located. A theory should be able to describe and explain how the properties of the wheel afford steering within this context.

*Identifying effective graphics.* For this research, I observed the evolution of graphics within SAE. For example, graphics of proofs and programming languages evolved to a text-based paradigm, whereas graphics of instructions for assembling furniture evolved to a pictorial paradigm. These observations helped me identify the types of graphics that might effectively
afford different types of actions. For example, text (proofs or programming languages) might be more effective than pictures for affording the communication of abstract concepts (such as mathematics), whereas pictures in instructions for assembling a piece of furniture might more effectively communicate how physical furniture parts should be assembled. I also focused on how afforded actions serve functional purposes within a given context.

Illustrative Example: Invention of the Pie Chart by William Playfair

The pie chart invented by William Playfair (1759–1823) was the original inspiration for the SAE approach employed in this dissertation, and serves to illustrate how this approach can inform the design of effective graphics. Spence (2006) described how Playfair invented the now-ubiquitous pie chart years before evidence from perceptual psychology and cognitive neuroscience could inform such a design. Amazingly, these charts have not significantly changed in the many years since their invention (Spence, 2006), suggesting that their properties effectively afford an author’s intention. How did Playfair design such effective graphic representations in the absence of a science for representation? Spence described how Playfair appropriated what appeared to be successful techniques used by cartographers, including the use of outlines and colour to indicate boundaries between geographic regions:

Natural selection, over the centuries, had ensured the evolution of effective maps and charts. Map features that worked well from a perceptual and cognitive aspect were retained, and other less successful constructions were discarded. Particularly during the 17th and 18th centuries, cartographers had eventually arrived at psychologically sound solutions to the problem of making maps and charts easy to read and use. Playfair imitated, adapted, and synthesized the lessons learned by mapmakers when he made his first statistical charts (Spence, 2006, p. 2433).

Inspired by this interpretation of Playfair’s strategy, I worked to identify examples that are likely to be ‘effective’ relative to an identifiable context (e.g., the communication of abstract concepts in the foregoing example) and functional purpose (e.g., to plot a safe path between landforms in an ocean in the foregoing example). This dissertation will explore how examining those ‘effective’ graphics, in relation to a context and functional purpose, can expose underlying principles. These principles, in turn, can serve as the basis for developing a theory.
SAE and Methodological Obstacles

By observing the evolution of graphics within SAE, we can avoid some methodological obstacles identified by researchers in the related human-computer interaction subfield of information visualization (Carpendale, 2006, 2008; Fekete et al., 2008). The non-laboratory settings in which information visualizations are most often used require many resources (in addition to visualizations) and often require extended periods of time. Visualization researchers find it difficult to observe and keep track of the wide variety of resources that audiences draw upon for given tasks, and over the extended periods of time during which visualizations are used (Carpendale, 2008). As a result, it is difficult for researchers to observe patterns that could inform theory development. Using SAE, we can avoid these difficulties by identifying environments where natural selection results in ‘effective’ types of graphics relative to the functional purpose that they afford. For example, in the Chinatown example, it seems reasonable to suppose that the photographs may inform actions of audiences from diverse linguistic backgrounds. The properties of the photograph can then be related to a functional purpose (such as communicating to audiences who read multiple languages), inviting theory development.

In the second half of this dissertation, I will use the famous controversy surrounding Euclid’s use of pictures in the *Elements*, and Proposition 35 of the *Elements* (Appendix I) in particular, as a site of artefact evolution in which types of graphics correspond to a clearly identified functional purpose, by examining under what conditions pictures may interfere with the communication of abstract concepts.

1.8 Outline

This dissertation’s model maps to the linear dissertation format as shown in Figure 1-4. Each column of Figure 1-4 focuses on different properties of graphics elucidated by the model. The ‘Pictorial vs. Symbolic Information’ column on the left of Figure 1-4 focuses on fundamental properties of graphics. The ‘Pictorially or Symbolically Represented Objects and Relations’ column on the right of Figure 1-4 is focuses on how objects and relations are pictorially or symbolically represented to produce composite pictures, diagrams, and sentences.
<table>
<thead>
<tr>
<th>Symbolic vs. Pictorial Information</th>
<th>Pictorially or Symbolically Represented Objects and Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background</strong></td>
<td><strong>Chapter 2</strong></td>
</tr>
<tr>
<td>Background pertaining to pictorial and symbolic properties of graphics to inform the model that enables claims in Chapter 3.</td>
<td>First part of Chapter 4</td>
</tr>
<tr>
<td></td>
<td>• Provides background for a linguistic account of graphic objects and spaces.</td>
</tr>
<tr>
<td></td>
<td>• A neuroscientific approach connecting the notion of graphic objects and spaces to the model initiated in Chapter 3.</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td><strong>Chapter 3</strong></td>
</tr>
<tr>
<td>Introduces a model of perception-reaction for conceptualizing:</td>
<td>Second part of Chapter 4</td>
</tr>
<tr>
<td>• <strong>Claim 1.</strong> Pictorial Information of a graphic and</td>
<td>• Extends the model to describe pictorially and symbolically represented objects and relations, enabling:</td>
</tr>
<tr>
<td>• <strong>Claim 2.</strong> Symbolic Information of a graphic.</td>
<td>• <strong>Claim 3.</strong> Composite Pictures as pictured relations among pictured objects.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Claim 4.</strong> Diagrams as pictured relations among symbolic objects.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Claim 5.</strong> Sentences as symbolic relations among symbolic objects.</td>
</tr>
<tr>
<td><strong>Predictions</strong></td>
<td><strong>First part of Chapter 5</strong></td>
</tr>
<tr>
<td>Predicts affordances of C1: Pictorial Information and C2: Symbolic information.</td>
<td>Second part of Chapter 5</td>
</tr>
<tr>
<td>• <strong>Prediction 1.</strong> Concrete Pictorial Prediction.</td>
<td>Predicts affordances of Claims 3-5 (composite pictures, diagrams, and sentences):</td>
</tr>
<tr>
<td>• <strong>Prediction 2.</strong> Abstract Symbolic Prediction.</td>
<td>• <strong>Prediction 4.</strong> Composite Picture Prediction.</td>
</tr>
<tr>
<td>• <strong>Prediction 3.</strong> Interference Prediction.</td>
<td>• <strong>Prediction 5.</strong> Diagrammatic Prediction.</td>
</tr>
<tr>
<td><strong>Demonstrations</strong></td>
<td>Chapters 6-7 demonstrate the implications of Claims 1-2 (pictorial and symbolic information) and Predictions 1-3 (pictorial, symbolic, and interference predictions; foreshadowed in the graphic below) by anticipating how pictorial and symbolic information afford the communication</td>
</tr>
</tbody>
</table>
of concrete or abstract concepts differently.

<table>
<thead>
<tr>
<th>Relations</th>
<th>Picture</th>
<th>Diagram</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictured</td>
<td>Pictured</td>
<td>Symbolized</td>
<td></td>
</tr>
</tbody>
</table>

Objects

- Pictured
- Symbolized

Intermediate

Early

Late

More Concrete

More Abstract

**Figure 1-4: How the dissertation is organized.**

In each column, following contextualization within relevant prior work (Figure 1-4, ‘Background’ row), a theoretical model is presented or extended (Figure 1-4, ‘Model’ row). The model is then employed to predict affordances of the graphics in question (Figure 1-4, ‘Predictions’ row). The predictions are then demonstrated by examining results of artefact evolution in a relevant site (Figure 1-4, ‘Demonstrations’ row).

Each row can be understood in relation to the scope and limitations described above: by synthesizing pre-scientific proto theories from the applied arts with scientific vernaculars from the cognitive neurosciences, each develops a way to describe properties of graphic representations at a Type I level (roughly ‘Background’ and ‘Model’ rows of Figure 1-4). Modelling seeks to develop at a Type II level, and ventures into a Type III theory for the purposes of illustrating the model developed thus far (‘Demonstration’ row Figure 1-4).

The model is presented in the linear parts and chapters as follows. Part I (‘Background,’ ‘Model,’ and ‘Predictions’ rows of Figure 1-4) introduces pictorial and symbolic information,
fundamental properties out of which different types of graphics are composed to varying degrees; Part II (‘Demonstrations’ row of Figure 1-4) builds on Part I by describing how pictures, diagrams, and sentences can be differentiated in terms of how objects and relations among objects are pictorially or symbolically represented.

**Part I: Model Development**

**Chapter 2: Model Development Part A: Reconciling Competing Accounts of Picture Perception**

Part I begins in Chapter 2 by introducing ideas that will be mined for theory development in Chapter 3 to conceptualize what it could mean for a graphic to be ‘pictorial’ or ‘symbolic.’ Following the theme of artefact evolution noted previously, Chapter 2 first reviews proto-theory from contemporary design and illustration and compares them to leading accounts from visual art theory, philosophy, and perceptual psychology, introducing a theoretical and practice-oriented landscape that will be drawn on in Chapter 3 to characterize perceptual capabilities that construct pictorial and symbolic properties of graphics in perception.

In particular, Chapter 2 compares constructivist (e.g., Goodman, 1968) and ecological accounts (Gibson, 1968; Kennedy, 1974) of picture perception to the practices of professional illustrators and designers, employing McCloud (1993) as a representative example. Each account is shown to have common properties with different visual processing streams of Goodale et al.’s (2005) dual-route hypothesis. Chapter 2 concludes by introducing a fundamental principal that explains the behaviour of the dual-processing streams of vision and that will serve as the foundation for the model introduced in Chapter 3: to describe how graphics make use of capabilities that enable humans to react, survive, and thrive in environments of dynamic change and variation.

**Chapter 3: Model Development Part B: Pictorial and Symbolic Properties of Graphics**

**Perception-Reaction Loop.** Chapter 3 draws on ideas synthesized in Chapter 2 to describe how graphics make use of capabilities that enable humans to react, survive, and thrive in the dynamic environments of everyday life.
Two technical terms are introduced in Chapter 3 to describe aspects of this perception-loop so that pictorial and symbolic properties of graphics can be distinguished and described: I use ‘emulation’ refers to that aspect of perception that ‘imitates’ the structure of the impinging light when that light structure fires sensory receptors of the retina. I use ‘simulation’ to refer to that aspect of perception that processes, filters, and interprets that which is emulated to predict/anticipate what is more distal, relative to the individual’s [impinging receptors], in space and time.

**Pictorial and Symbolic Properties.** Chapter 3 employs the model of perception-reaction noted above to describe pictorial properties as the emulated properties of a graphic. Additionally, Chapter 3 describes symbolic properties in terms of how the emulated properties from a graphic are processed and filtered through memory structures to simulate an author’s intended [objects, relations, or attributes] meaning. This notion of an author’s simulated intended meaning, and how that meaning can be recognized (or perceived) via a graphic, will be developed in order to describe how graphically represented environments are distinguishable from ‘actual’ or ‘non-represented’ environments.

**Chapter 4: Distinguishing Diagrams, Pictures, and Sentences**

Chapter 4 extends Chapter 3’s model of pictorial and symbolic information to create the foundation that will enable it to distinguish **pictures**, **diagrams**, and **sentences** in terms of how *objects* or *relations* are pictorially or symbolically represented. This approach to distinguishing types of graphics will require the development of a working notion for what it could mean for an *object*, *relation*, or *attribute* to be graphically represented. Chapter 4’s efforts first build upon prior work identified in the literature by examining how practitioners have distinguished types of graphics, such as diagrams and sentences:

Engelhardt’s (2002) theory of graphic objects and spaces is an application of linguistics to graphics. Engelhardt’s application of linguistics to graphics would ordinarily be deemed incompatible with ecological accounts of picture perception, which rejects picture-as-language theories (from Chapter 2). However, Chapter 2’s synthesis of ecological and constructivist approaches, through Goodale’s (2005) dual route hypothesis, provides an avenue for linking Chapter 3’s version of pictorial and symbolic properties to Engelhardt’s linguistically-based theory of graphic objects and spaces. This linkage between Chapter 3 and Engelhardt (2002) is
further facilitated by drawing upon Hurford’s (2003) application of the dual-route hypothesis to formal logic and written language: Hurford attributes the properties and roles performed by objects and predicates of formal logic to the architecture of the dual processing streams reviewed in Chapter 3.

Hurford’s interpretation of the dual-route hypothesis (introduced in Chapter 2) is employed to extend Chapter 3’s model of perception and reaction to conceptualize objects as targets for homeostatic reactions that are in spatial-temporal relations among other potential target objects that are distinguished by attributes, such as shape, texture, or color. In this simplified model, attributes enable the simulation or enactment of appropriate reactions to target objects, such as appropriate reaching, grasping, or gripping actions.

Chapter 4 reviews how authors can make use of the above to conceptualize: Diagrams as composed of pictorially represented relations among symbolically represented objects, sentences as composed of symbolically represented relations among symbolically represented objects, and pictures as composed of pictorially represented relations among pictorially represented objects.

Chapter 5: Predicted Affordances of Graphics

Chapter 5 further develops pictorial and symbolic information in order to predict affordances of each by synthesizing Norwich’s (1991) recruitment of Shannon and Weaver’s (1959) information theory with more recent findings focused on the development of perceptual and conceptual categories (Mandler, 2006) to describe how information emerges in perception, and is shaped by the interaction with inherited capabilities and experiences of a life-span.

Part II

The focus of this dissertation’s second part is to demonstrate the model introduced in the dissertation’s first part by applying the model to a specific case study, thus enabling the scientific vernacular developed in that part to be demonstrated and refined. Starting in Chapter 6, the chosen case study is a famous controversy surrounding Euclid’s use of pictures in the Elements, and Proposition 35 of the Elements in particular (Appendix I). Euclid’s use of picture proofs, and the long history of controversy surrounding those picture proofs, provide a rich set of clues that,
when explored via the model, demonstrate properties and affordances of pictures relative to symbols, and sentences in particular.
Part I: The Model
2 Pictorial vs. Symbolic Part A: Competing Accounts of Picture Perception from the Literature

2.1 Preface

The previous chapter introduced the main project, discussed the requirements for the project’s main objective (developing a theoretical model), described an approach to making observations (using SAE), and introduced an example of artefact evolution (a simplified pattern that the model should explain, i.e., the axiomatic paradigm, using the case study of Proposition of Euclid’s *Elements*). This chapter will introduce ideas that can be applied to develop one part of the proposed model. Specifically, it will focus on the idea of ‘representation,’ which can be thought of in various ways: cognitive theorists have conceptualized it as occurring ‘in the head,’ while the applied visual arts and design have traditionally focused on the craft of configuring materials in the world without considering internal processes of the mind. This chapter will go on to explore how these views compete, which will set the stage for theory development in the next chapter.

2.2 Introduction

Scholars working in the fields of the fine and applied visual arts define picture perception differently from those working in perceptual psychology. Within the arts, the human ability to perceive pictured objects is characterized as learned, or constructed, like a ‘visual language’ (Gombrich, 1960; Goodman, 1976; Kulvicki, 2010). Within perceptual psychology, picture perception is often characterized as an unlearned, innate ability; according to one dominant account, this unlearned biologically inherited capability to perceive actual surfaces and edges means that only optical properties of light produced by pictures are required to perceive ‘pictured’ surfaces and edges (Gibson, 1978; Gibson, 1971; Kennedy, 1974; Lee et al., 1980; Juricevic et al., 2006; Hammad et al., 2008). This chapter will demonstrate how these competing views can be reconciled by applying Goodale et al.’s (2005) dual route hypothesis. To date, no previous research has considered the ramifications of this hypothesis with respect to the arts or the philosophy of art.\(^5\)

\(^5\) This was the case as of November 2010 (see Kulvicki, 2010a).
**Thesis of this chapter.** According to current research in cognitive neuroscience, visual processing occurs in two ‘streams,’ defined by where the processing generally takes place in the brain: a *dorsal* stream and a *ventral* stream (Goodale et al., 2005). Visual processing in the ventral stream produces and uses traces (‘memories’) of previously visually-processed objects to inform the selection of potential actions, such as a particular kind of grasping action (Goodale et al., 2005), using memory systems similar to those used for language (Martin et al., 2001). Once a particular action is selected, visual processing in the dorsal stream relies less on memory and instead directs the selected action in relation to target objects that are visually processed in real time (Goodale et al., 2005).

Much of the discussion that follows will focus on demonstrating how the competing claims from the fields of art theory and perceptual psychology correspond to descriptions of these streams. I will describe how the ventral ‘what/how’ stream’s memory systems enable actual and depicted objects to be taken as symbolized information, underpinning the constructivist account used in the arts. Next, I will describe how the real-time visual processes of the dorsal stream correspond to the optical account from perceptual psychology, where processing of actual and pictured features relies less on prior learning and memory. I will also demonstrate how pictorial information engages these processes more than the ventral processes.

![Figure 2-1: Competing theories of this chapter presented on a spectrum.](image-url)
**Chapter scope and limits.** This chapter will explore the unresolved debate between proponents of ecological and constructivist accounts, by focusing on a rare exchange between scholars in the fields of arts and sciences. This debate was sparked by Ernst Gombrich, a noted art historian, and J.J. Gibson, and will be highlighted in the next section.

**Key terms.** An ‘ecological account’ is a theory according to which picture perception is independent from constructivist processes (learned social-cultural conventions).6 ‘Dual-processing account’ refers to Goodale et al.’s (2005) hypothesis, which is biologically-based, but includes a role for prior experience (i.e., conventions) in perception. Competing theories of this chapter are presented on a spectrum in Figure 2-1, with the ‘ecological account’ to the left and the ‘constructivist’ to the right. ‘Innately learned,’ in the middle, refers to processes that are learned, but that could be learned in similar ways across populations that develop in environments with similar properties.

**Chapter outline.** McCloud (1993) developed a theoretical ‘comic book’ to clarify the types of graphic representations used by the authors of comic books. Professional illustrators, in fields as diverse as biomedical illustration and human-computer interaction, have relied heavily on McCloud’s work (see e.g., Zhang, 1997; Mathis, 2010; Manning, 1998). The following sections will describe McCloud’s work and then explore the more philosophical domain of visual art theory, specifically research conducted by Nelson Goodman (1976), who built on Gombrich’s work (1960) to conceive graphic representation as a visual language. As noted in Chapter 1, this led to a major debate between Goodman (a philosopher of aesthetics) and James J. Gibson (a perceptual psychologist) about whether picture perception is a learned or unlearned capability. Gibson was a pioneer in the field of the psychology of perception; his research included work about picture perception where he took a non-linguistic approach to graphic representation (see e.g., Gibson 1978). The present discussion herein is related to the domain of cognitive neuroscience, a field that is closely related to perceptual psychology and may help resolve the debate.

---

6 ‘Constructivist processes’ could be thought of, in the context of this dissertation, as akin to ‘learned processes.’
2.3 Context

In 1960, Ernst Gombrich, an art historian, published a highly influential book called *Art and Illusion: A Study in the Psychology of Pictorial Representation*. Soon after, Gibson (1960) proposed a refutation of Gombrich’s ‘constructivist’ account and presented his own ecological account. A debate ensued, first involving Gibson and Gombrich, and later including Rudolf Arnheim (Gombrich et al., 1971), and Nelson Goodman. Goodman later developed his own constructivist account, and the constructivist account became solidified within art theory, philosophy of art, and other branches of philosophy. Gibson’s ecological account is still used in the field of perceptual psychology (e.g., Kennedy, 1974; Lee, 1980; Hammad et al., 2008) and has inspired affordances-based ideas of design introduced more recently (e.g., Norman, 2002).

In contrast, professional illustrators working in fields as diverse as biomedical illustration and human-computer interaction commonly employ McCloud’s (1993) theoretical ‘comic book’ about graphics to inform and explain design decisions. McCloud’s work is one example of what I call a proto-theory used in applied practice. However, it introduces ideas that are akin to those that are contested by other theorists from outside of the art and craft of design, inviting comparison between the ideas of those theorists. In particular, it is helpful to compare constructivist and ecological accounts of picture perception to McCloud’s account, because leading ecological and constructivist theorists participated in the rare debate between the fields of visual art theory and perceptual psychology described above: the Gibson-Gombrich debates. This chapter will compare aspects of the debate with aspects of McCloud’s work, and show how these can be used to develop concepts that help explain pictorial and symbolic properties and affordances of graphics.

2.4 An Applied Account: McCloud

McCloud (1993) arranged graphic representations into a triangular space, similar to those shown in Figure 2-2, to define what it means for a graphic representation to be ‘realistic,’ ‘iconic,’ and ‘linguistic.’

The left of the figure represents ‘reality,’ and successive graphic representations are arranged on a ‘dimension’ to the right of the figure. Each successive graphic representation has a less isomorphic relation to the ‘real’ object, becoming more ‘iconic.’ McCloud defined ‘realism’
as complex, objective, and specific, and defined ‘iconic’ as simple, subjective, and universal. As the representations extend further to the right, the objects become increasingly ‘linguistic,’ involving ‘meaning’ in terms of words and sentential structures.

Figure 2-2: A modified version of McCloud’s (1993) triangle.

‘Received’ vs. ‘Perceived’ Information. McCloud identified two kinds of information produced by graphic representations. He defined pictures as producing ‘received’ information, arguing that audiences require no prior learning to ‘get the message’ from received information, so the message is ‘instantaneous.’ In contrast, he defined written language as producing ‘perceived’ information, arguing that it takes time and specialized knowledge to decode the abstract symbols of language. He went on to argue that when pictures are ‘abstracted from reality,’ they require a higher level of perception, like words.

Discussion: McCloud’s Contentious Claim

According to McCloud, learning plays a role in the ability to distinguish pictorial and symbolized information: picture perception is a less-learned capability and symbol recognition is a more-learned capability. His claim that picture perception is an unlearned capability is controversial, especially as it relates to visual art theory and the science of perception.
2.5 An Ecological Account of Picture Perception

According to Gibson, pictures make use of an organism’s biological capabilities to perceive ‘visual information.’ This sub-section reviews his theory of visual information and direct perception, followed by his theory of picture perception (Gibson, 1978).

Visual perception. In Gibson’s theory, light reflects from environmental surface features to observation points. The light at a particular observation point has a unique structure, called an optic array, which corresponds to surface features.\(^7\) An optic array at one position differs from one in another position (and is the basis for disparity). Some properties of the array, but not all, change between observation points. Gibson referred to non-changing properties as ‘invariants,’ as in the phenomenon of an object looming larger at an accelerating rate as an object is approached (see e.g., Lee & Kalmus, 1980). Gibson described the invariants in optic arrays as ‘visual information’ (Gibson, 1978).

Picture perception. Expanding on Gibson’s (1971) research, Kennedy (1974) described a representational picture as one layout of surfaces that can be artificially treated and arranged to provide visual information for or about a different environmental layout.\(^8\) Kennedy claimed that by inspecting these surfaces, observers can identify things that are not present: “That which is present in the optic array from the picture is a frozen, perhaps exaggerated, moment in the set of transformations that would reveal the invariant” (1974, p.45). Kennedy supported his claim that this was an unlearned capability based on case studies from previous research. For example, Hochberg and Brooks (1962) observed a child who had been raised in an environment devoid of pictures such as photographs and outline drawings. The child was later able to perceive and recognize objects represented in outline drawings and photographs without being trained to do so, supporting the theory that humans can perceive and recognize represented objects without being taught a ‘visual language.’ Based on this and other evidence, perceptual psychologists tend

\(^7\) Clouds, for example, generate an optic array, but to some it might not seem that the light corresponds to any ‘surface’ (it has more to do with the total cloud density between the observer and the sun or sky). I think that a Gibsonian would argue that ‘surfaces’ are merely collections of particles. In this sense, clouds have a ‘surface’ that can reflect light.

\(^8\) An interesting question arises here regarding whether or not Kennedy’s (1974) definition applies to written sentences, such as text, because text involves some artificially arranged optical invariants that do not change based on observational perspective. This issue will be addressed in the next chapter.
to accept the ecological account. However, theorists working in other fields, especially the visual arts, often reject this ecological account of picture perception, particularly after arts theorists increasingly rejected positivist accounts, and perhaps because at the time biological accounts could not account for the numerous cultural phenomena that artists were using in their works.

**Discussing Gibson**

A Gibsonian would likely agree that a member of a culture unfamiliar with pictures could recognize a line drawing of a piece of fruit, such as an apple; a Gibsonian would likely also agree that this individual would not recognize a particular (learned) meaning from outside his/her culture in a line drawing (e.g., the outline drawing of an apple that denotes Apple Corporation). The Gibsonian account of picture perception involves an implicit constructivist capability; this is the key to integrating the two theories. The next section will explore this constructivist account in more detail.

**2.6 A Constructivist Account of Picture Perception**

Expanding on Gombrich’s research, Goodman (1976) presented an account of picture perception that also differed from Gibson’s constructivist account. Goodman argued that almost any picture may represent almost any other thing, and that pictures, and their relationships to objects, are parts of a constructed system of representation.

**Pictures as labels.** Goodman argued that pictures are labels, like linguistic predicates with contextually-relative meaning: “just as a red light says ‘stop’ on the highway and ‘port’ at sea, so the same stimulus gives rise to different experiences under different condition” (Goodman, 1968, p.19).

**Syntactic density.** According to Goodman, pictorial ‘labels’ can be distinguished from other denotation systems, such as language, by their syntactic density (Goodman, 1976).

**Critique of resemblance.** Goodman did not deny that pictures can resemble their referents, but argued that any object can resemble any referent in some way, and therefore a learned conceptual category is required to focus on meaningful correlations within a particular social or cultural context. For example, all objects contain molecules, and therefore resemble each other. Goodman wrote, “there is no innocent eye […]. Not only how but what [the eye] sees
is regulated by need and prejudice. [The eye] selects, rejects, organizes, associates, classifies, analyzes, and constructs” (Goodman, 1976, p. 7–8). For Goodman, even rules of perspective and ‘realism’ are arbitrary and conventionally established (Giovannelli, 2010, citing Goodman, 1984 and Goodman, 1976). Thus, for Goodman, ‘realistic’ pictures are those that are depicted using a familiar system of correlation. In other words, the observer always needs a learned ‘key’ to read a picture.

**Discussing Goodman**

Does Goodman’s constructivist account completely rule out the role of innate biological processes (of the ecological account)? It does not seem to, because learning relies on innate biological processes. Even if the ability to perceive an object is constructive, observers must still use their eyes to access light rays. Certainly, Goodman would never claim that eyes are anything other than genetically inherited biological structures, even if conventions somehow govern the use of those eyes and other processes involved in perception. Thus, Goodman’s constructivist account may permit an implicit role for innate biology. The issue of contention may be more related to how picture perception relies on conventions that make use of innate biological processes (see Fig. 2.1).

**2.7 Discussing Ecological, Constructivist, and Applied Accounts**

McCloud’s view that pictures are ‘received’ is similar in some ways to the view held by ecological psychologists: picture perception is an unlearned capability. The ecological psychologists did not investigate graphics that are clearly symbolic, but McCloud discussed symbolic graphics, and the obvious role played by learning in developing capabilities to use symbolic graphics. His view supported aspects of both ecological constructivist accounts; indeed it is difficult to imagine anyone would claim that reading is an unlearned capability.

When discussing symbols, McCloud alluded to how the distinction between pictures and symbols is not clear-cut, and how pictures can have symbol-like properties: “When pictures are more abstracted from ‘reality’, they require greater levels of perception, more like words” (McCloud, 1993, p.153). He went on to describe the pictorial properties of written words: “When words are bolder, more direct, they require lower levels of perception [and] are received faster, more like pictures” (1993, p.153). Thus, although he distinguished between pictures and
symbols, he also acknowledged the symbol-like properties of pictures and the picture-like properties of symbols, foreshadowing the model being developed in this dissertation: *pictorial and symbolic properties are found in all graphic representations, but to varying degrees.*

2.8 A Path to Reconciliation via the Dual-Route Hypothesis

One way to identify points of integration between these views is to explore when, where, and how perceptual capabilities develop on top of, or in conjunction with, inherited (biologically grounded) capabilities. For example, an ecological psychologist is unlikely to deny that capabilities to use symbols are learned (most people remember learning to read). Additionally, a constructivist is unlikely to deny that an individual’s eyeballs are biologically inherited. In my previous research (Coppin, 2011), I reconciled the ecological and constructivist views of picture perception by employing Goodale et al.’s (2005) dual-route hypothesis; the next section describes this reconciliation.

2.8.1 Dual Route Hypothesis

Independent of the debates between Gibson and the art theorists reviewed earlier in this chapter, Goodale et al. (2005) proposed a hypothesis that can help integrate the competing accounts (see Coppin, 2011). Goodale et al. identified two distinct but interrelated perceptual-cognitive processes: ‘vision-for-perception’ and ‘vision-for-action.’ He began by arguing that evolution must have been driven by the need to direct movements in response to environmental changes, and that perception was one step in that process. Complex cognitive operations (such as recognition, identification, and planning) enabled humans to predict future outcomes and thereby be capable of even more flexible and adaptive behaviours.

**Distinct functions of the two streams.** According to Goodale et al. (2005), the dorsal stream enables real-time control of motor actions. In contrast, the ventral stream yields the rich and detailed representation of the world required for cognitive operations such as recognition, identification, and prediction. For example, the dorsal stream may allow an individual to reach out and grasp objects, but it is “trapped in the present” (Goodale et al., 2005: 4): it can deal only with objects that are visible when the action is being planned or ‘programmed.’

**Lesion studies.** Lesion studies research can help clarify the roles played by these systems. Patients with lesions in the dorsal stream can find it difficult to use vision to form a
grasping or aiming movement toward objects that appear outside the range of foveal vision, a deficit often described as optic ataxia (Goodale et al., 2005 citing Bálint, 1909). An opposite pattern of deficits and visual abilities has been reported among patients with visual form agnosia, where brain damage is concentrated in the ventral stream (Goodale et al., 2005, citing Perenin and Vighetto, 1988). For example, with visual form agnosia, an individual can touch and reach things fine, but cannot tell what they are, in the sense of what category they belong to—e.g., toy or furniture, etc.

**Discussing the Dual Route Hypothesis**

Goodale et al.’s (2005) account of the dorsal ‘where’ stream describes how it processes visual qualities about the environment in ways that are less influenced by learning (as demonstrated by the patients with object agnosia). In this way, the dorsal processing stream is consistent with the ecological account. Goodale et al.’s (2005) account of the ‘what/how’ stream describes how connections to ventrally located memory structures facilitate ‘recognition’ and prediction. According to this account, the ‘what/how’ stream enables perceivers to link a present moment to previous memories, allowing them to plan possible actions (as demonstrated by patients with optic ataxia; see Fig. 2.2). In this way, the ventral processing stream is consistent with the constructivist account.

**2.9 Synergistic Aspects of Constructivist and Ecological Accounts**

Let us now consider how constructivist and ecological accounts need to be considered not mutually exclusive; at some level, each rely upon independent but interrelated neural systems that enable survival in dynamic environments. As discussed above, both ecological and constructivist accounts seem to be required, so they may index psychological processes that occur in different regions of the brain.

The main issue here is where picture perception relies on capabilities that build on, or use, processes that rely less on learning for perception. The dual-route hypothesis has explanatory power here, because it integrates perceptual-cognitive processes that require more learning with those that require less learning.
For some direct forms of action, visual processing seems to be less influenced by learning, supporting the ecological account. Actions guided by visual processing of pictured surfaces and edges are possible without learned conventions, but recognition requires learned conventions. For example, patients with object agnosia have memory deficits in regions connected via the ventral stream; many can correctly engage in (e.g., grasping) actions in relation to actual and pictured surfaces and edges, but are unable to recognize those surfaces and edges as particular objects (Goodale et al., 2005). This finding supports the ecological account in which learned conventions are not necessary to ‘perceive objects’ in pictures (Kennedy, 1974; Gibson, 1978). Thus, learned conventions are not always required for visual processing related to action, but they are probably always required for recognition.

Recognition requires learning and memories, corresponding to the constructivist-recognition account. The ability to recognize an object (i.e., visually processing light rays conveying visual information about an object and making this meaningful by connecting it to a personal narrative of past experiences and possible future actions) arises from the ventral stream’s interaction with other multi-modal anterior and lateral regions. The ventral stream codes complex visual information, thereby supporting visual representation of objects bereft of meaning. Categories emerge through experiences, and solidify through repeated experiences (Goodale et al., 2005). This process shares some properties with Goodman’s constructivist account of ‘pictures as labels’ and the role played by a conceptual model in enabling a particular object to reference another object or idea (Goodman, 1976).

Goodman’s (1976) critique of resemblance applies to recognition, but possibly not to visual processing. He claimed that any feature of an object or picture could be considered to resemble any other object or picture, so a learning process is required to train a perceiver to recognize the key features of an object that can be interpreted as referring to other objects. This process shares some features with how memory helps create a hierarchical schema in the ventral stream (e.g., Moscovitch, 1992; Goodale, 2005).

Goodman’s (1976) claim that pictures are linguistic, and that words-as-language and pictures-as-language differ in terms of their degree of syntactic denseness, is supported by

---

9. The next chapter (section 3.3) provides a more detailed description of ‘representation.’
evidence that the ventral processing stream connects to memory structures that are also associated with language. A growing body of evidence is indicating that language and object recognition both involve ventrally connected memory systems (Martin et al., 2001). A language-like role in object recognition corresponds to Goodman’s (1976) picture-as-name account.10 So, what is the key difference between pictures and spoken/written language? Goodman (1976) argued that pictures and spoken language are both conventionally based, but that pictures are more ‘syntactically dense’ than words. A Gibsonian would likely draw a sharp distinction between language and visual perception, countering Goodman’s ‘visual language’ account. In contrast, according to the dual-route hypothesis, two processing streams of the same object occur simultaneously, but one has ‘language-like’ attributes. This idea of ‘language-like attributes’ will be important to the discussion of symbolic information in the next chapter, and should help shed some light on the perplexing issue introduced at the beginning of this chapter: are icons pictures or symbols? The simple answer is that icons are both; the next chapter will show how pictorial and symbolic information are components of all graphic representations, but to varying degrees. The ventral processing of visually perceived objects can be understood as language-like and syntactically dense (see Goodale et al., 2005, expanding on Goodman, 1976).

What is at issue? At this point, the constructivist account appears to involve a link between a perceptually processed picture and traces of prior perceptions. The ecological account appears to involve the visual processing of pictures, but not recognition, and appears to involve neural systems for perception-for-action.

2.10 A Neuroscientific Account of ‘Recognition’

This section focuses on some distinctions between (learned) recognition-for-prediction and (less-learned) visual processing-for-action, to help clarify the mechanisms that enable conventionalization. Drawing from Barsalou (2009), it will present a simplified description of encoding and retrieval to provide a biological conception of how the emergence of learned categories enables recognition. Within this discussion, conceptual categories act as the basis for conventions used in pictures, including pictures that are processed as symbols.

10 Extending on Barsalou’s (2009) research, the idea here is that memories are inherently situated, multi-modal, and stored as traces in the same systems used for perception. Concepts and categories follow the same situated and multi-modal pattern. From this perspective, visual language is possible.
Perception (encoding). When features (e.g., of Apple 1) are perceived, detectors fire for edges, surfaces, colour, etc. Conjunctive neurons in association areas capture the activation pattern. Populations of conjunctive neurons code the pattern, with each individual neuron participating in the coding of many different patterns. Local association areas near a sensory modality capture activation patterns within it (e.g., the visual features of Apple 1). These consolidations can be considered ‘memory traces.’

Cohesion and consolidation. Higher association areas in temporal, parietal, and frontal lobes integrate across modalities, particularly when conscious attention is focused on a feature. This architecture has the functional ability to produce modal re-enactments (simulations).

Recognition (retrieval). Now suppose Apple 2 appears. Once again, feature detectors fire, but memory traces of Apple 1 are used for Apple 2. Because a trace is linked to other aspects of the prior experience, those prior experiences are reconstructed from the traces, perhaps consciously (simulated). This recognition, recollection, or remembering produces additional traces, and together the process is the basis for learned conceptual categories; the emergence of categories enables the symbolization of information through repeated perception and recognition.11

2.10.1 Path to Reconciliation

This section summarizes how ecological and constructivist accounts may align with two distinct, but interrelated, visual processing streams.

Perception-for-action enables an organism to react to changes in real time. Organisms that can simulate possible actions (predictions) can prepare reactions to future changes, thus enabling survival in dynamic environments.12 One aspect of learning is an organism’s capability

---

11 For example: suppose I have driven for many years. Am I visually perceiving-for-prediction or am I visually processing-for-action? ‘Obeying’ the red light requires a capability to predict the outcome of not obeying the red light. However, the repeated encoding and cohesion-consolidation of stopping at red lights in everyday life ‘automates’ this process, turning it into a ‘lower-level’ non-conscious process, which Vygotsky (1978) referred to as ‘fossilization.’ At this stage, the process seems (and probably is) more like perception-for-action.

12 Some might claim that prediction is not ‘simulating possible action,’ e.g., predicting your own death does not require that you simulate (reenact) dying. However, Barsalou (2009) would probably argue that the concept of dying emerges from past experiences that cause the concept of death to emerge, even if those experiences are mediated.
to construct predictions from traces of past perceptions (‘memories’), and can be considered the basis for conceptual categories. According to Barsalou:

An important issue concerns how simulators for abstract concepts develop. Barsalou (1999) proposed that simulators for abstract concepts generally capture complex multi-modal simulations of temporally extended events, with simulation of introspections being central. According to this account, simulators develop to represent categories of internal experience just as they develop to represent categories of external experience. (Barsalou, 2009, p. 1282)

Recognition-for-prediction enables an individual to retrieve that which has been learned, in order to form a prediction about a (future) reaction. Generally, these processes fall along two separate but interrelated streams. Dorsal processes are related to an individual’s action targets in real time, while ventral processes are related to the mechanisms that enable recognition, retrieval, and prediction.

Recognizing visually processed features as objects requires using traces from previously processed features. Thus, ventrally-connected memory systems are implicated. Graphic designers use this recognition capability to create representations that refer to traces from the past: ‘symbolic information.’ In other words, when an individual visually processes a graphic representation, certain aspects of the processed information are meaningful because of their prior experiences.

**Back to pictures.** What this all means is that humans can react to visually processed actual or pictured features without being trained to do so, supporting the ecological account. Thus, dorsal ‘trapped in the moment’ processes are implicated. Graphic designers use this ability to communicate via pictures: because pictures involve little reference to traces from the past, the optical structure of the visual information carries the majority of the intended meaning. This meaningful optical structure can be termed ‘pictorial information.’ However, some pictures may through language. For example, the concept of ‘infinity’ could be introduced in, for example, an educational institution. Later use of this concept (e.g., a homework assignment for a physics class) would be a reenactment of the initial experience of being introduced to the concept (and the simulations conjured during that introduction, because a physics lesson would also reference the student’s prior experiences). As the concept is reused in everyday life, it leaves new traces, causing the concept to become ‘more abstracted’ from the original classroom/textbook experience. This process is fundamental to the discussion of symbolic information in the next chapter.
trigger re-enactments/simulations, for example avoiding fire, or avoiding running a finger along a sharp edge. In the proposed model, every graphic representation contains pictorial and symbolic information, but to varying degrees. For example, when someone visually processes the optical structure produced by a graphic representation, the aspect of the visually processed information that is meaningful independently of prior experience is ‘pictorial information.’ When the properties of this pictorial information enable retrieval of simulations constructed from learned conceptual categories and conventions, this is ‘symbolized information.’

2.11 Conclusions

This chapter reviewed three competing accounts of picture perception, ultimately identifying a foundational principle that could serve as a means to integrate and reconcile each competing account, and serve as a foundation upon which this dissertation's model could be based: the perception-reaction loop. To identify this foundational principal, this chapter applied a leading hypothesis from the field of cognitive neuroscience to show how two visual processing streams can correspond with competing accounts of picture perception.

The discussion of artefact evolution in Chapter 1 suggested that some proto-theories might foreshadow some aspects of the questions at issue, because applied practitioners who produce graphics for audiences must survive or thrive within a competitive marketplace (e.g., professional illustrators and designers common use of McCloud’s (1993) theoretical ‘comic book’). In this way, they are applying a proto-theory in practice. However, McCloud’s model introduced ideas that are similar to those contested by other theorists from outside the field of design, inviting comparison among ideas. Because leading theorists of ecological and constructivist accounts participated in a rare debate between the fields of visual art theory and perceptual psychology, this chapter compared constructivist and ecological accounts of picture perception to McCloud’s version. By comparing McCloud’s ideas with aspects of the Gibson-Gombrich debate, this discussion helped develop concepts that will subsequently be useful to explain pictorial and symbolic properties and affordances of graphics.

Next, the chapter then incorporated some of my previous research (Coppin, 2011), specifically an example of how developments in cognitive science/cognitive neuroscience can inform debates within the fields of art-design-media, by showing how competing accounts correspond to two aspects of an interrelated neural process that enables the human perception-
reaction loop: the so-called dual processes of vision. In this previous research, I demonstrated how competing claims from art theory and perceptual psychology correspond to descriptions of these streams. I argued that the ventral ‘what/how’ stream’s memory systems enable actual and depicted objects to be taken as symbolic, underpinning the constructivist account used in the arts, whereas the real-time visual processes of the dorsal stream correspond to the optical account from perceptual psychology, in which processing of actual and pictured features relies less on prior learning and memory. I also demonstrated how pictorial information engages dorsal processes more than ventral processes.

Taking advantage of the fact that many other neural systems could be used for the same (sort of) explanation, the next chapter will explore more fundamental principles underlying perception-reaction, which are at the core of the proposed model. This is the idea of a homeostatic perception-reaction loop, a process that enables an organism to maintain conditions necessary for its survival and prosperity in the face of environmental change and variation.
3 Pictorial vs. Symbolic Part B: Theory Development

Preface

The previous chapter demonstrated how empirically-based findings from the neuro- and cognitive sciences – extending Goodale et al.’s (2005) work to graphics – could potentially help reconcile accounts of picture perception that were, previously, seemingly at odds with one another. This chapter will focus on conceptualizing more fundamental principles that could liberate the proposed model from the need to conceptualize the properties of graphics in terms of this specific neural system. These more fundamental principles were foreshadowed in the last chapter, with the suggestion that Gibson and Goodman might have been referring to two different aspects of an interrelated process, but disagreed because no vernacular was available to decompose it. For example, they both used the word ‘perception,’ but Goodman was referring to perception-with-recognition of an object’s type or class, whereas Gibson was referring to a simpler recognition-free component of perception. Let us now consider how the relative engagements of these two aspects can help distinguish between the symbolic and pictorial properties of graphics.

3.1 Introduction: Perception-Reaction to Environmental Change and Variation

This section introduces and demonstrates a simplified model of perception-reaction to help develop a model to distinguish between the pictorial and symbolic properties of graphics. This model was inspired by a key point made by Goodale (2005), which was reviewed in the previous chapter: “evolution must have been driven by the need to direct movements in response to environmental changes, and that perception was one step in that process.” Complex cognitive operations (such as recognition, identification, and planning) enabled humans to predict future outcomes and thereby be capable of even more flexible and adaptive behaviors” (see Chapter 2).

---

13 This approach has common properties with the notion of autopoiesis described by Varela, Thompson, and Rosch, (1999) and cybernetics (Wiener, 1948).
The model developed in this dissertation considers how graphics use (and perhaps even extend) this ‘perception-reaction loop’: To get a sense of the simplified perception-reaction model, imagine an apple is sitting on the table in front of you. You reach out, grasp it, start to take a bite, realize it is rotten, and then toss it into the trashcan across the room.

Two interrelated processes enable this perception-reaction-loop. First, light rays were reflected from the apple and picked up by retinal detectors in your eyes, and then processed by your perception-reaction system, enabling your body to reach for and grasp the apple. Second, grasping the apple required positioning your arm and hand (in three-dimensional space) such that your fingers could curl around to the distal part of the apple not observed (‘sensed’) by retinal detectors. Your body was able to ‘predict,’ ‘anticipate,’ ‘fill in,’ or ‘simulate’ the shape of the non-visible side of the apple with sufficient certainty to form your hand and fingers into a configuration that enabled successful grasping of both the detected and not-detected parts of the apple.

This chapter will introduce two terms to denote two interrelated aspects of this perception-reaction process, which will subsequently help us distinguish between the pictorial and symbolic properties of graphics: \textit{emulation} and \textit{simulation}. The relative engagements of these two aspects will help distinguish between the symbolic and pictorial properties of graphics: I will describe how pictorial properties engage relatively more ‘emulation,’ while symbolic properties engage relatively more ‘simulation.’ I will demonstrate that the relevant aspect of emulation is the \textit{structure} of the light that is perceptually processed from the physical environment (or graphic). I will also demonstrate that the relevant aspect of simulation is how the item being emulated is processed to enable the viewer to appropriately anticipate, predict, or 'simulate' reactions to potential environmental change and variation. By applying the previous chapter's neuroscientific account of perceptual processing (see Barsalou, 2009), it is possible to develop a more complete picture of emulation and simulation.

\footnote{Another benefit of beginning with the idea of perception-reaction to environmental change and variation is that the stage is set for a conceptualization of pictorial and symbolic information of a graphic, where information is treated as a measure of uncertainty, following Shannon’s seminal theory.}

\footnote{These two aspects are based on evidence from the cognitive sciences. They are described using various labels; my reasons for using these particular terms are discussed below.}
When I perceive and act by grasping an apple, my body is using memory trace resources to form the shape of my hand (via simulation) to meet the unseen totality of the shape of the apple.

- These memory trace resources are the remnants of prior perception-reactions, in which detectors had fired for edges, surfaces, and colour, etc. Conjunctive neurons in association areas captured the activation pattern. Populations of conjunctive neurons coded the pattern, with each individual neuron participating in the coding of many different patterns. Local association areas near a sensory modality captured activation patterns within it (e.g., the visual features of an apple) (Barsalou, 2009).

- Thus, memory trace resources include neural connections among lower-level association areas. These lower-level association areas are themselves neural connections among even lower-level association areas. The lowest levels of the hierarchy are more tightly coupled to sensory detectors of specific perceptual modes such as vision, sound, and touch. At the lowest levels of the network are sensory receptors that are triggered by sensations/stimuli (e.g., retinal detectors of the human eye) (Barsalou, 2009).

- Higher-level association areas of the network (in the temporal, parietal, and frontal lobes) integrate across modalities such as vision, sound, and touch. This architecture has the functional ability to produce modal re-enactments (simulations) to engender reactions to environmental change and variation (Barsalou, 2009).

Thus, between the processes of perception and reaction, what begins as a modally specific sensation engenders less modally specific (more amodal) simulations of possible actions, ultimately resulting in motor responses to specific targets of an environment.

**a. Simulation.** Inspired by Barsalou's (2009) use of the word, here 'simulation' refers to the aspect of perception-reaction that processes, filters, and interprets the item that is picked up by detectors to predict/anticipate or 'simulate' what is more distal in space and time. This aspect of perception allows us to see distal objects as three-dimensional objects, even when only some subsets of two-dimensional surfaces reflect light to our eyes. To achieve this, our visual systems must be able to ‘simulate’ things and events in the world, in some spatio-temporal sense.
Researchers have demonstrated that this process relies on experience/memory and learning (Hockema, 2004; Anstis, Verstraten & Mather, 1998; Goldstone, 1998; Freyd, 1992).

**b. Emulation.** Here, the word ‘emulation’ denotes the aspect of perception-reaction that is most closely coupled with the proximal stimuli and sensations that impinge upon a viewer. In contrast with simulation, which is constructed from memory traces of past perception-reactions, 'emulation' refers to the structure of that which is picked up or detected from the environment. It connotes how the detection of proximal stimuli and sensation ‘imitates’ the structure of proximal stimuli that is detected. With respect to vision, emulation would involve the near ‘isomorphic’ response of retinal receptors to the aspects of the optic array (see Gibson, 1978) to which they are specifically tuned to respond.\(^{16}\) Because the item being emulated is processed further from this ‘surface interface’ and closer to higher-level aspects of the perceptual-cognitive system, it becomes increasingly less accurate to characterize this process as emulation (and more accurate to describe the process as simulation). The key characteristic of this aspect of perception-reaction is the structural relationship between the viewer’s response and the proximal stimulus (change) to which the response is a reaction.

One way to distinguish simulation from emulation in practice is that simulated aspects of perception are more malleable than emulated aspects. Consequently, the range of possibilities for a simulation that is a response/reaction to an external change/variation is greater than for emulation. Additionally, because structural correspondence (between the proximal stimuli/change and the reaction) is not a defining characteristic of simulation (unlike emulation), it is not easy or even possible to ‘map’ the reaction directly back to the structure of the stimuli. In sum, the key characteristic of simulation is (subjective) extrapolation from the proximal structure of stimuli to the recognition of the distal structure of the environment.\(^{17}\)

---

\(^{16}\) However, the retinotopic maps in V1 (the primary visual cortex), which bear a topographic relationship to the light stimuli impinging on the retina, while not ‘isomorphic’ with the stimuli, would still qualify as highly emulated.

\(^{17}\) The distinction I am making here between emulation and simulation can be compared to the distinction classically made between sensation and perception, but with a different emphasis because, as will be explained in subsequent chapters, I align these processes with pictorial and symbolic properties and affordances of the representations/information, as well as more recent results in cognitive science (e.g., Goodele et al., 2005 and Barsalou, 2009). My use of the word ‘emulation’ is akin to classic notions of sensation, but emphasizes how the structure of stimuli picked up from an environment performs a role that engenders perception of physical structures, akin to Gibson’s (1986) optic array; this emphasis will provide a means to describe pictorial properties of graphics.
Consider Figure 3-1, which presents a picture and drawings/diagrams of an apple and the Apple logo. The emulation that occurs when you perceptually process the drawings of the apple is structurally similar to the emulation that occurs when you perceptually process a real apple. In contrast, the activity that occurs when you perceptually process the Apple logo (and consider Apple Corporation and its various products, for example) is less easily mapped back onto proximal stimuli, and is therefore classified more as simulation. Additionally, different (and more) learning is required to perceive-recognize the Apple logo than is required to perceive-recognize a drawing of an apple. Together, these two perceptual-cognitive distinctions justify distinguishing between the two graphics such that one (on the left) is classified as ‘more pictorial’ and the other (on the right) as ‘more symbolic.’

![Figure 3-1: Pictures vs. symbols.](image)

Other examples can help demonstrate how the words ‘pictorial’ and ‘symbolic’ are conceptualized in terms of emulation and simulation. If I see a picture of food (e.g., the Chinatown example in Chapter 1; Figure 3-1, left), light reflects from the printed surface to my point of observation where it is emulated by my perceptual system. If I am reading written text or printed characters (Figure 3-1, right), light reflects from the page to my point of observation where the light is also emulated by my perceptual system. In both cases, pictorial properties of a graphic are emulated. However, other symbolic properties emerge during perceptual processing, when the emulated pictorial properties are processed and filtered to draw on memory traces from prior perception-reactions to construct simulations of an item or property that is beyond (or other

(Section 3.3.3). My use of the word simulation is synergistic with a growing body of work that describes the predictive capabilities of perception (e.g., Norwich, 1981; Hockema, 2004; Bubic, 2010; Clark, 2013); this emphasis will provide a means to suggest how an audience predicts (‘simulates’) what an author intended to symbolically represent (Section 3.3.4).
than) what is emulated from the marked surfaces of the graphic. The *failure* to read text written in an unfamiliar language helps distinguish symbolic from pictorial properties more explicitly. In the Chinatown example, I could emulate (and physically point to or trace the outlines of) the written graphics, but because I could not read the foreign language, I lacked the memory trace resources required to simulate the symbolic properties of the graphic.

Having introduced the concepts of pictorial and symbolic in terms of emulation and simulation, it is now important to consider how we distinguish graphically represented items from non-graphically represented items in an environment. The next subsections will expand on the concepts introduced above, and show how perceiving or recognizing an author’s intention plays a role in the construction of a graphic’s properties during perceptual processing. First, I will demonstrate how this conceptualization of graphics is linked with how information (associated with an environment or graphic) is a function of, and emerges during, perceptual processing.

### 3.2 Information (of a Graphic or Environment) as a Function of Perception

To this point, I have only used the concept of *information* implicitly. As stated in the introduction, ‘information’ here refers to the classic definition introduced by Shannon and Weaver (1948), who conceptualized information as a reduction of uncertainty. The next chapter will present a more technical account of information and apply it in the proposed model. However, before developing a way to distinguish graphically represented items from non-graphically represented items, it is important to show how the model conceptualizes information of a graphic (or environment) as a function of (and not independently of) the perceptual mechanisms and capabilities that process it.

---

18 This view of information will be linked to Shannon’s theory of information (Shannon, 1948; Weaver, 1959) in the next chapter, which is based on probabilistic models interpreted as involving uncertainty. The range of possibilities for simulation is analogous to the amount of mathematical uncertainty in Shannon’s theory. Norwich (1985) united Shannon’s probabilistic theory with perceptual theory, but this discussion is beyond the scope of this subsection’s focus on distinguishing graphically represented items from non-graphically represented items in an environment.
This section uses a strategy similar to that used in the previous section by conceptualizing information as a function of perception, beginning with three key assumptions regarding how perception links an organism to its environment:

- **Assumption A:** An organism is something living (‘surviving’) in an environment and needs to maintain the conditions of its own viability despite changes in its environment.
- **Assumption B:** Perception is ‘designed’ (via natural selection) to enable reaction to spatially and temporally variegated environments.
- **Assumption C:** Perceptual systems are naturally selected.

Within this context, organisms can respond to perceptually processed changes in their environment by taking actions that change themselves (and possibly the environment) in ways that maintain their viability and their survival, at least to the point of reproductive success.19

Understanding the environment in which we function can provide important insights into how our perceptual systems work, and the particular aspects of the environment to which they are attuned. To illustrate with a metaphor, suppose that a cast-bronze statue is lost, but its mould is not. Observers could still describe many properties of the lost statue, because some of the features of the statue correspond to the mould that shaped it. Similarly, the properties of a changing environment (‘the mould’) can reveal some characteristics of an organism’s perceptual system (‘the statue’) – specifically, those ‘moulded’ by the forces of natural selection. Based on the above assumptions:

- The principles of natural selection imply that a perceptual system that enables an organism’s survival would be ‘designed’ (via natural selection) to react to the properties of a changing/varying environment.
- It follows that the dynamically varying properties of an environment that are relevant to the organism’s survival must be perceivable.
- The relationship between environmental properties and capabilities means that evolutionarily-relevant properties of a dynamically varying environment can be used to identify the capabilities of an organism’s perceptual system.

---

19 Here, ‘environment’ means the area surrounding the organism, which can (directly or indirectly) causally impinge on the organism.
Thus, expanding on Gibson (1978), I characterize a *perception-reaction loop* as the process that enables a homeostatic organism to react to changes in its immediate environment and thereby maintain its viability in that environment. Information makes this process possible: ‘information’ (as I am using the term) is a proximal change that corresponds to a distal change/variation in an environment, such that an organism can perceive a distal change/variation in the environment. In this way, information can be considered to emerge at the intersection of an organism’s perceptual capabilities and its environment.

### 3.3 Graphics as Intentional Visual Information

The previous two sections explored the difference between pictorial and symbolic properties in terms of their involvement in two perceptual-cognitive processes: ‘emulation’ and ‘simulation.’ The statue/mould example illustrated how perceptual-cognitive aspects are only one element of these processes; less emphasis has been placed on the environment that engages those processes (i.e., *external* graphics, like those studied in arts/design schools). It is important to develop a more balanced strategy in which the concepts of ‘pictorial’ and ‘symbolic’ can yield a ‘unified’ understanding of graphics that incorporates both perceptual-cognitive and media-based aspects. This section explores how graphics can be related to the perceptual-cognitive processes described above.

Fields such as art, philosophy, and semiotics have a long history of research about graphics (graphic representations). A full analysis of this history is beyond the scope of this chapter, but rather this section summarizes the properties of external representations that engage with the psychological processes described above. Briefly, ‘graphic’ can be defined as: *an environment that is intentionally configured to produce information that causes an intended visual percept.*

For example, most scholars would agree that the text-sentences on this page are examples of graphic representations. This section will demonstrate that when readers believe they have understood the meaning of a text-sentence, to some extent they believe that they have perceived/understood an author’s intent. Guided by this vague idea of ‘perceiving an author’s intent,’ I will introduce another example – not of text-sentences, but of an environment that is

---

20 Gibson (1978) used the term ‘perception-action loop.’ While perception is inextricably tied to action (*reaction*), certain aspects of a perception-reaction loop are ‘other’ than reaction; these aspects cannot be completely separated.
intentionally configured by an ‘author’ to cause an intended perception in an observer. In this example, I will decompose a series of simple situations to develop a model for graphics that can apply to both pictures and symbols in terms of perceived intention. Section 3.3.1 and Figure 3-2 will distinguish between aspects of a perceived environment to yield a narrower representation of perception. The illustrations that follow include:

- Perceiving-recognizing random environmental changes relative to another organism’s (re)action(s) (see Section 3.3.1; Figures 3-2a and 3-2b).

- Perceiving-recognizing an organism’s reaction relative to another organism’s intentional (re)action(s) (see Section 3.3.2; Figures 3-2b and 3-2c).

- Perceiving-recognizing another organism’s (re)action, which was intended to cause an intended (visual) perception (i.e., a (graphic) representation) (see Section 3.3.3; Figures 3-2c and 3-2d).
Figure 3-2: a) An environment, b) an environment that has been reacted to, c) an intentionally reacted to (configured) environment, d) an environment intentionally configured to cause an intended perception.
3.3.1 Perceiving-Recognizing Random Environmental Changes Relative to an Organism’s Reaction: A Face Statue in a Forest

Imagine you are in a forest. You see a stone artefact that appears to be carved into the shape of a face (Figure 3-2d). For illustrative purposes, let us stipulate that this carved stone face is an example of an external graphic representation.\(^{21}\) Using the vernacular developed thus far, both the carved face (an external graphic representation) and the natural environment that is not the carved face (Figure 3-2a) are perceived via information. What property distinguishes the information related to the external graphic representation from the information related to the non-face environment?

The non-face environment can be recognized, based on learning and memory traces from past perceptions, as the product of natural environmental changes (and entropy). The viewer recognizes the carved face as having been produced by an organism (the environment has been modified by an organism’s reaction). However, defining an external graphic representation as ‘an environment that is perceived-recognized as having been acted on or reacted to by an organism’ is insufficient to distinguish ‘information related to a representation’ from ‘information related to a non-representation’ (e.g., random environmental change-variation). For example, suppose you see animal tracks in the mud beneath the statue (Figure 3-2b). Even if you were to perceive-recognize these tracks as having been caused by an organism, intuitively you would likely not perceive-recognize the organism’s ‘intent.’ Even if the track was a by-product of an organism, such as a person, engaging in an ‘intentional’ (re)action, the organism was likely barely ‘aware’ of where and how the tracks were being produced.

**An intentional reaction that is not intended to cause an intended percept.** Suppose you perceive-recognize a rusty can lying on the side of the trail (Figure 3-2c). Intuitively, it seems that the can was ‘intentionally’ created for a purpose, but it was probably not placed there to convey something meaningful to you or anyone else. To distinguish between the carved statue and the discarded can, the next section demonstrates how the discarded can might be perceived

\(^{21}\) Gibson and Goodman might disagree about whether the ability to perceive this representation is learned or unlearned, but they would both probably agree that it is an external representation.
as having been intentionally configured to cause an intended perception, and is therefore an ‘external representation.’

3.3.2 Distinguishing an Intentional Reaction from a Reaction that is Intended to Cause an Intended Perception

The following example involves two secret agents who must communicate by using external representations that are visible only to themselves, and not to others.

Suppose two nations are at war, and one nation is planning to invade the other. Invader Nation wants to invade Defender Nation, but a key harbour in Defender Nation may or may not be filled with explosive land mines. Invader Nation sends two secret agents into Defender Nation to determine whether there are mines in the harbour, to inform the invasion decision. The first agent (Agent Sender) will observe the harbour and communicate observations to a second agent (Agent Receiver), who will return to Invader Nation and report back to the military leaders there. Agent Sender must leave messages that are visible to Agent Receiver but not to the citizens of Defender Nation.

Both agents agree to the following ‘conventions’ (codes) in advance: at a particular crossroad on a particular path, a rusty can on a particular tree stump is meant to indicate Situation X (“the harbor is filled with deadly mines”). A rock on the tree stump is meant to indicate Situation Y (“the harbor is not filled with deadly mines”). An empty tree stump indicates that Agent Sender has not left a message.

Agent Receiver walks the trail each day and inspects the tree stump. After months of no message (empty stump), Agent Receiver sees a rusty can on the stump, and interprets this to mean that Agent Sender has indicated Situation X ("the harbor is filled with deadly mines"). As hoped, none of the other travelers on the trail (citizens of Defender Nation) perceive the rusty can as anything other than random garbage (Figures 3-2a, b, c).

Using key terms of the model introduced thus far, the citizens of Defender Nation all would probably perceive the random trees, leaves, cans, etc., in the woods as the result of natural environmental changes (Figures 3-2a). They might perceive the can or rock on the stump to be an
organism’s (re)action (Figures 3-2b and 3-2c). However, only the agents would perceive the message, because of their previously agreed upon code-convention; they perceive the can on the stump to be more than an organism’s (re)action. In other words, Agent Receiver perceives Agent Sender’s (probable) intention when perceiving the rusty can on the stump – and the rusty can on the stump represents Situation X.22

What distinguishes a can as ‘garbage’ from a can as ‘representation’? The agents developed simulations together that enable them to process an author’s intentions perceptually, using information from an environment that is intentionally configured (by the author) to cause an intended perception. This is an external representation.

A. Agent Sender can simulate how Agent Receiver will perceive Agent Sender’s reaction-action of placing the can.

a. Agent Sender’s ability to simulate how Agent Receiver will perceive Agent Sender’s placement of the can had an origin. Through prior interactions with Agent Receiver, Agent Sender had learned how Agent Receiver would perceive the can (where Agent Receiver’s perception of the can follows the scenario outlined in B2). Agent Sender had learned how Agent Receiver would perceive Agent Sender’s intent. Simultaneously,

B. Agent Receiver can simulate how Agent Sender will perceive Agent Receiver’s action-reaction of the placed can.

b. Agent Receiver’s ability to simulate how Agent Sender will perceive Agent Receiver’s perception of the can had an origin. Through prior interactions with Agent Sender, Agent Receiver had learned how Agent Sender would perceive the can when placing it, which means Agent Receiver had learned how Agent Sender would simulate Agent Receiver’s perception of the can (where Agent Sender’s simulation of Agent Receiver’s perception of the can follows the scenario described in A1). Agent Receiver had learned how Agent Sender would intend the can to be perceived by Agent Receiver.

---

22 Can an intention be perceived? Recall that perception can be characterized as consisting of emulation and simulation, where emulation is the sensed visual information (e.g., from faces of a three-dimensional object exposed to a perceiver), and simulation fills in the rest of the ‘object’ (e.g., the non-exposed surfaces not facing the perceiver). Because the intentions of an author are simulated, intentions can be perceived.
SUMMARY (A): When Agent Sender intends to create information (via a graphic), this means that Agent Sender can simulate how Agent Receiver will perceive via this information.

SUMMARY (B): When Agent Receiver perceives via Agent Sender’s intended information, this means that Agent Receiver can simulate how Agent Sender intended Agent Receiver to perceive via the information.

Figure 3-3: When Agent Sender sends a signal (1), Agent Sender can simulate (2) how Agent Receiver will perceive the signal (3). When Agent Receiver receives a signal (4), Agent Receiver can simulate (5) how Agent Sender simulated (6) how Agent Receiver would perceive the signal (7).
Here, ‘intention’ means:

- an Organism A’s Simulation Z
  - of how Organism A’s reaction
  - will be (perceived and) Simulated Y by Organism B,
    - in that A’s Simulation Z of
      - B’s (perception) Simulation Y
        - includes a simulation of B’s (perception) Simulation Y
          - of A’s Simulation Z.

---

**Figure 3-4: Sharing perspectives through external representation.**

Thus, information is ‘intended’ – *a representation is information intentionally created to cause an intended perception.*
3.3.3 Pictorial Information

This section focuses on a challenging problem: pictorial information. As discussed in Chapter 2, Gibson and his supporters argued that picture perception is an unlearned and/or innate capability, and supported their argument with empirical evidence. At first glance, this might suggest that the concept of graphics and external representation discussed in Section 3.3.2 might not apply to pictures, because organisms would have no way to learn/develop a ‘shared understanding.’ This section will describe two interrelated kinds of learning that can serve this purpose: ‘perceptual learning’ and ‘cultural learning.’ It will show how perceptual learning corresponds to Gibson’s ‘ecological account’ and how cultural learning corresponds to Goodman’s ‘learned-conventionalized’ account.

3.3.3.1 Perceptual Learning

Recall from Section 3.2 that information can be conceptualized as a function of perception, so each organism can be conceptualized as living in a distinct ‘perceptual world.’ Now consider a scenario in which organisms survive and/or thrive in a common environment with other organisms, and consider that perceptual emulation can be conceptualized as isomorphic to (the sensation of) the perceived environment. Here, individualized perceptual worlds could be constructed from identical, similar, or overlapping sources: a common environment. This section will discuss how this common environment can enable perceptual simulations with common properties among organisms, how these common properties can enable a ‘shared understanding,’ and how the authors of external graphic representations can use this shared understanding to communicate.

If two organisms perceptually emulate identical, similar, or overlapping aspects of the same environment using similar perceptual systems, then the simulations they later construct using traces of these perceptual emulations may share common properties that could serve as the basis for a kind of shared understanding – one that authors of graphics can use to communicate. Although the simulations may not be identical, the space for possible simulation types should converge over time, assuming similar neurobiological perceptual systems within organisms of the same species operating within a shared environment. This is especially true if each organism is able to adapt to its environment via perceptual learning (see Gibson, 1979), such that
perceptual emulations become more isomorphic to the proximal stimuli (information) about similar relevant distal objects or events. In this kind of situation, it is difficult to distinguish between what is biologically inherited (innate) from what is learned. The next section will apply this idea of perceptual learning to pictures.

3.3.3.2 CLAIM 1: Pictorial Information

Section 3.1b explained how environments can be emulated. Section 3.1a clarified how emulation leads to learned capabilities that can simulate environmental features. Authors of graphic representations can use these capabilities by producing visual information that causes intended emulations and simulations, thus enabling perceptual processing of environmental features (surfaces and edges) that are not part of the material/configured surface. In other words, pictorial information can be defined as:

1. visual information (Section 3.2),
2. from a graphic (Section 3.3),
3. that causes
   (a) emulation (Section 3.1b) and thus also
   (b) learned simulation (Section 3.1a) of
       1. environmental surfaces
       2. that are not part of the material (e.g., configured/marked) surface.

For example, visual information from a painting of an apple produces pictorial information. What does it mean to perceive a surface that is not actually part of the environment being perceived (see Kennedy, 1974)? An evolutionary account has explanatory value here. Humans perceived their environments for millennia before the first picture-making, and picture perception is based on this inherited biological capability. When a perceiver is presented with visual information, the perceptual system behaves based on how evolutionary forces selected it to behave, whether the visual information is from a ‘natural’ environment or is artificially
produced via a representation (typically a marked surface). He/she perceives a ‘surface,’ even if the ‘surface’ is not part of the marked surface.\(^{23}\)

### 3.3.4 Symbolic Information

#### 3.3.4.1 Cultural Learning

This section introduces a scenario that could speed the process of shared intentionality construction, which I will later use to develop a simple model of ‘communication’ in which graphic representations are used to conceptualize symbolic information. Recall the account of perceptual learning in which (conspecific) organisms in a shared environment typically develop simulations with common properties, which authors of graphic representations can use to communicate (Section 3.3.4.1). In addition to allowing organisms to perceive ‘objects’ in their environment, such simulations enable perception of other organisms in the environment and may include aspects of their behaviour. To do this well, an organism would also need to develop the reflexive ability to simulate what other organisms can perceive. In an environment with many other conspecifics, when this ability leads to a shared understanding of intentions, this process can be defined as ‘cultural learning’ (e.g., Tomasello et al., 1993). The next sections show how this process can help distinguish between pictorial and symbolic information.

#### 3.3.4.2 CLAIM 2: Symbolic Graphic Information

Authors can configure an environment to produce pictorial information (see Section 3.3.3.2); when this information is perceptually processed, simulations are formed from the memory traces created during perceptual learning. Similarly, authors can configure an environment to produce pictorial information; when this information is perceptually processed, simulations are formed from memory traces created during cultural learning. This is ‘symbolic information.’

\(^{23}\) See Hammad et al., (2008) for an example of how difficult it is for a human perceiver not to perceive pictured surfaces and edges.
Symbolic pictorial (graphic) information is:

1. Pictorial information (Section 3.3.3.2),
2. which, when emulated, uses
   i. memory traces produced during cultural learning (Section 3.3.4.1), to
   ii. simulate an environment (Section 3.1a) that is beyond (or other than) what
       is pictured.

See Fig. 3.1 for an example of symbolic pictorial information: the logo of Apple Inc. Most observers (see Kennedy, 1974) would perceive the logo as a pictured apple. However, cultural learning allows observers to recognize the pictured apple as referring to Apple Inc., which is not pictured. In other words, even though the concept of Apple Inc. is not pictured (e.g., its products, physical location, market share, etc.), mental simulations of Apple Inc. as a concept are produced in the observer when he/she recognizes the apple as the company logo. The pictured apple ‘symbolizes’ Apple Inc.

3.4 Conclusions

Chapter 2 illustrated that constructivist and ecological accounts do not have to be mutually exclusive, using Goodale et al.’s dual-route hypothesis. However, it may be difficult to distinguish between pictures and symbols (or pictorial and symbolic information) in terms of neural systems, because other neural systems may also be involved. Therefore, Chapter 3 demonstrated how pictures and symbols (and pictorial and symbolic information) can be distinguished using principles that transcend particular neural systems.

This account rests on the homeostatic notion of a perception-reaction loop: an organism must maintain the conditions necessary for its survival (and prosperity) by perceiving and reacting to environmental changes. This perception-reaction loop to environmental change involves two interrelated aspects: emulation and simulation. Emulation is similar to sensation, and structurally emulates the local changes in the organism’s environment. Simulation is triggered during emulation and is less structurally isomorphic to the sensory signal, allowing the organism to simulate (possible) variations and structures not directly perceived (e.g., the far side of a piece of fruit that a hand wraps around during a grasping action).
Pictorial information can be conceptualized as an author’s intentionally emulated (visual) information from a graphic representation; symbolic information can be conceptualized as an author’s intentional simulation, constructed when emulating a graphic representation. This version of pictorial and symbolic information rests on a conceptualization of representation that is characterized in terms of intention: representation is an environment (e.g., a marked surface) that is intentionally configured by an author to cause an intended emulation (and consequently an intended simulation) in the audience.

This conceptualization also rests on the idea of ‘shared intentions,’ which refers to the capability of an audience to simulate an author’s intent. When organisms with common biologically capabilities develop in an environment with common properties (especially when they can communicate), they can develop conventions.

**Moving Forward**

This chapter developed the beginnings of a Type I theory (as introduced in Chapter 1). To be classified as Type II or III, a theory needs to explain and predict. Before beginning this task, it will be necessary to discuss another type, or class, of graphic representation that at first glance seem to confound the pictorial versus symbolic distinction developed thus far: hybrid graphics representations (e.g., diagrams). The next chapter focuses on finding a way to distinguish pictures from diagrams and diagrams from sentences.
4  Pictured vs. Symbolic Objects and Relations: Diagrams vs. Sentences

4.1 Introduction

The previous chapter distinguished between pictorial and symbolic information and pictorial and symbolic graphic representations. This chapter focuses on graphic representations that seem to be hybrids of pictorial and symbolic graphic representations, such as diagrams.

![Diagram](image)

Figure 4-1: An example of a diagram, which appears to have pictorial and symbolic properties.

The diagram presented in Figure 4-1 includes ‘pictorial’ elements, such as spatial relations between the three vertically arranged squares. It also contains canonical symbols: the letters “C,” “A,” and “B.” The next sections will develop a model for hybrid graphic representations, which can be used to predict their affordances. First, it is important to review how both ‘visual’ graphic representations (e.g., Engelhardt, 2002) and graphically represented text-sentences (e.g., Hurford, 2003) have common properties: objects, object attributes, and relations among objects. I claim that these common properties across modes are not arbitrary; they emerge because graphic representations (symbolic or pictorial) use a common visual cortex (see Chapter 2; Goodale et al., 2005, cf. Hurford, 2003). The human visual cortex is often characterized as a multi-modal, task-oriented system used for emulation and simulation (Barsalou, 2009; Ehrsson et al., 2003; Kosslyn et al., 2001; Lotze et al., 1999; Porro et al., 1996). The spatial cortex’s task-oriented system is involved in visual perception-recognition, but also in language generation and comprehension (Glenberg & Robertson, 2000; Bergen & Chang, 2005). From this perspective, language-implicated simulations are re-enactments of prior perceptual
emulations and use much of the same perceptual-cognitive machinery used during perceptual emulation. This can help clarify how the common neural system and processes within the spatial cortex enable objects, relations, and attributes to emerge in speech, writing, and visual graphic representations. Extending the discussion about pictorial and symbolic information in Chapter 3, this chapter will explore how pictorially or symbolically representing object relations and/or attributes may support or interfere with mental emulation and simulation of object relations and/or attributes. This will help develop a model to distinguish between various graphic representation types (such as realistic pictures, outline drawings, diagrams, and sentences; see Table 1, Chapter 1) based on whether object relations, shapes, or details are pictured or symbolic. Within this model:

- realistic pictures are considered to be pictured relations among pictured object shapes and shape details;
- outline drawings are considered to be pictured relations among pictured shapes, but without pictured details;
- diagrammatic graphics are considered to be pictured relations among symbolic graphic objects; and
- sentential graphics are considered to be symbolic relations among symbolic objects.

This model will enable preliminary hypotheses to be framed about about the perceptual-cognitive affordances of graphic representation types typically used in design. Combining the theoretical categories set out in the model with findings from neuroscience and cognitive science related to the capabilities of the spatial cortex, attention, and working memory, (for example that emulation and simulation share common, limited resources) it will follow that:

- pictured object relations or attributes can interfere with (an author-intended) mental simulation(s) of object relations or attributes;
- pictured object relations or attributes can scaffold, or support (an author-intended) mental emulation;
- combinations of pictured and symbolic information can free resources for mental simulation of symbolic objects; and
- symbolic information affords mental simulations that are difficult, or impossible, to emulate.

Outline. The following sections will conceptually bridge previous research to develop a model for graphic objects, relations, and attributes, and show how these can be pictured or symbolic. Following the established strategy of SAE, the discussion will begin with a review of proto-theories from practitioners, followed by theoretical work.

Neural correlates for visual language. The first bridge is between Engelhardt’s 2002 account of graphic objects and spaces and the conception of pictorial and symbolic information
presented in Chapter 3. Continuing with the strategy introduced in Chapter 1 – using artefact evolution to inform literature reviews – I will begin with Engelhardt (2002) because this work synthesized the proto-theories of leading graphic representation designers (visual information designers). Next, I will extend Hurford’s (2003) argument about the neural basis for predicate-argument structure (this also builds on Goodale et al.’s dual-route hypothesis). According to Hurford, “neural evidence exists for predicate-argument structure as the core of phylogenetically and ontogenetically primitive (prelinguistic) mental representations” (2003, p. 261). Hurford argued that the structures of modern natural languages can be mapped onto neural processes and structures. I will extend Hurford’s argument to map the properties of visual ‘languages’ (e.g., see Engelhardt 2002) onto what Hurford referred to as primitive ‘representations,’ and thereby provide a possible neural basis for Engelhardt’s theory of graphic objects and spaces.24

4.2 Prior Work

4.2.1 Engelhardt: A Synthesis of Visual Information Design

Engelhardt (2002) synthesized research about graphic representation, drawing from practitioners including Bertin (1983; a cartographer who wrote a seminal work about graphic theory), Tufte (2006, 1990, 1997, 1983; a statistician turned information designer), and Card et al. (1999; computer scientists from the field of human-computer interaction and information visualization). Engelhardt built upon these prior works to develop a comprehensive theory of graphic representation from a linguistic, semantic, and syntactic perspective. The next section summarizes part of Engelhardt’s theory, specifically his view of graphic objects and spaces.

4.2.1.1 Graphic Objects and Spaces

Engelhardt applied principles from linguistics to graphic representation, describing pages and panels as units (similar to words) with visual elements and sub-elements (similar to morphemes). He treated arrows and other connectors as various parts of speech, and described how graphic representations perform many semiotic roles. The following aspects of his theory are relevant here (see Figure 4-2):

24 Here, 'representation' refers only to external (graphic) representations (e.g., pictures or symbols), not internal mental representations.
• A graphic space (e.g., timelines or geographic coordinate systems) is composed of graphic objects.
• Graphic objects are symbols or pictures, e.g., dots, pictograms, or arrows.
• The objects relate to each other spatially, e.g., through nodes, connectors, frames, or labels.

Graphic spaces can also be graphic objects, and graphic objects can be graphic spaces, enabling complex hierarchies, not unlike language. Additionally, Engelhardt described text in sentences as a special case of graphic representation; this point will be important to the following discussion of diagrams relative to sentences.

Figure 4-2: Graphic objects and spaces.

4.2.2 Neural Basis for Graphic Objects and Spaces: Hurford

In his 2003 article, Hurford argued that the predicate-argument structure of first-order predicate logic has a neural basis, adapted to a sentient organism’s traffic with the world, rather than having to be postulated as ‘logically true’ or even ‘Platonically given.’ He claimed that neuroscience can yield informative answers to the question of where elements of logical form came from. This subsection will demonstrate how Engelhardt’s conceptualization of graphic objects and spaces can also be interpreted as involving a neural basis – in the visual cortex and streams as described by the dual-route hypothesis – and that common properties emerge across modes because common neural systems are implicated in perception across these modes.

Hurford illustrated his theory using PREDICATE(x) notation. ‘PREDICATE’ refers to attributes of an object, such as ‘red,’ ‘round,’ ‘big,’ or ‘small,’ whereas the variable in parentheses,
‘(x),’ represents the object that the predicates describe or are assigned to. By linking predicates to a common variable, as in:

\[ \exists x [\text{ROUND}(x) \& \text{GREEN}(x) \& \text{APPLE}(x)] , \]

multiple predicates can be represented as belonging to a common object ‘x.’ The logical statement above roughly translates as: “There is a round green apple.”

The main point of Hurford’s account is that the dorsal ‘where’ stream (identified in Chapter 2 as ‘Gibson-like’) identifies where objects are, but not what objects are:

One process is the rapid delivery by the senses (visual and/or auditory) of information about the egocentric spatial location of a referent object relative to the body, represented in the parietal cortex. The eyes, often the head and body, and sometimes also the hands, are oriented to the referent object, which becomes the instantiation of a mental variable. (Hurford, 2003, p. 273, emphasis added)

According to Hurford, this idea of a ‘mental variable’ is critical because in his synthesized account, the dorsal ‘where’ stream cannot identify what objects are. The ventral ‘what/how’ stream (described in Chapter 2 as ‘Goodman-like’) assigns attributes to the mental variable, enabling the identification/recognition of objects. Hurford wrote:

The other process is the somewhat slower analysis of the delivered referent object by the perceptual (visual or auditory) recognition subsystems in terms of its properties. The asymmetric relationship between the predicate and the variable, inherent in the bracketing of the formula, also holds of the two neural processes. (Hurford, 2003, p. 273)

In this theory, the two modes of processing are very different and perform different roles. The ‘where’ process is implicated less in identification, and the ‘what’ process includes an intermediate stage wherein objects are ‘dynamically localized’ (Bridgeman et al., 1994).

Hurford continued with a discussion of the instantiation of the mental variable (‘FOPL’ is an acronym for first-order predicate logic’):
The story so far, then, is that the brain interprets relatively abrupt discontinuities—such as change of orientation of a line, change of colour, change of brightness—together as constructing holistic visual objects which are expected to share a “common fate.” It is these whole objects that are held in attention. A shift of attention from one object to another is costly, whereas a shift of attention from one feature of an object to another feature of the same object is less costly. This is consistent with the view underlying FOPL, that the entities to which predicates apply are objects, and neither properties nor locations. In accepting this correlation between logic and neuropsychology we have, paradoxically, to abandon an “objective” view of objects. No perceptible physical object is ever the same from one moment of its existence to the next. Everything changes. Objects are merely slow events. What we perceive as objects is entirely dependent on the speed our brains work at. An object is anything that naturally attracts and holds our attention. But objects are what classical logicians have had in mind as the basic entities populating their postulated universes. The tradition goes back at least to Aristotle, with his “primary substances” (individual physical objects). (Hurford, 2003, p. 276)

I will be mining these ideas in the next sections to explore the neural basis for graphic objects and spaces. First, it will be helpful to present an illustrative example. Figure 4-3c presents a picture of an apple (shape, texture, colour). Figure 4-3a graphically presents a symbolic representation of the apple using first-order predicate logic, and Figure 4-3b presents its English translation. According to Hurford’s theory, the mental mechanisms that underlie our perception of the apple as a green object are also the mental mechanisms that make logical predication of the sort shown in panel A possible.
There is a round green apple

**Figure 4-3:** A round green apple graphically represented via a picture, logic, and an English sentence.

At this point, I can extend the analyses presented in Chapters 3 and 4 to explore the emergence of objects, relations, and details in spoken, written, and visual languages.

## 4.3 Pictured Objects

### 4.3.1 Two Ways of Explaining

This subsection explores two synergistic explanations for pictured objects. The first explanation extends the conceptualization of symbolic *information* provided in Chapter 3 to include an account of symbolic *objects*. The goal is to describe how authors can configure a representation to use an audience’s capabilities to emulate objects of a natural environment, and thereby enable emulation of (represented) ‘objects’ that are *not part of* the configured surface. Here, pictured objects are similar to ‘illusions’ of natural objects. The benefit of this first explanation is that it does not require a formal definition for ‘objects’ (or for ‘attributes’ or ‘relations’). Individuals are somehow able to emulate objects of a ‘natural’ environment through some unknown mechanism or process; a growing body of evidence indicates that objects, attributes, and relations emerge across different modes because each of these ‘languages’ uses this unknown, but common, neural mechanism or process (cf. Pulvermüller, 2005; Glenberg,
2008). This first explanation is more conservative than the second explanation, but it also requires dealing with the idea of ‘objects’ that are emulated but ‘not there.’

According to the second explanation, pictured objects are actual, but intentionally authored, objects; they radically extend an individual’s homeostatic perception-reaction loop by increasing the space for possible reactions in space-time, using external representations to increase the space for resources (simulations, memory traces, and emulations) from which reactions are formed. By increasing the ‘width’ and ‘breadth’ of the homeostatic perception-reaction loop, an individual’s capability to survive and thrive relative to competing organisms is radically increased.

In this second explanation, objects are targets for homeostatic reactions, and graphically represented objects increase the space for possible homeostatic reactions to include a wider space for change-variation. An explanation borrowed from telerobotics can help develop a simplified model for a graphic representationally afforded perception-reaction loop, which may reveal how this homeostatic loop is afforded by external representations in repositories, archives, public policy communication, education, and science.

4.3.2 Explanation #1: Pictorially Represented ‘Objects’ Make Use of Capabilities to Perceive Non-Represented Objects of an Environment

Recall the pictorial information discussed in Chapter 3. By using an audience’s naturally-selected capabilities to emulate actual environments, an author can intentionally cause emulation of an ‘environment’ that is not part of the configured surface (e.g., markings on piece of paper) using a bottom-up sensation from a graphic representation. For example, a common explanation for the ability of humans to see in colour is that it helped our ancestors to differentiate red objects (berries) from green backgrounds (e.g., Regan, 2001). A contemporary designer can use this capability for a non-berry detection task. For example, traffic stop signs are typically red, possibly because this makes them more easily seen against a background of trees or a blue sky.

---

25 This process also radically increases a society or culture’s perception-reaction loop; for the sake of simplicity, this discussion is restricted to individuals.
Similarly (although from a functional perspective), human emulation capabilities might not have been naturally selected to enable us to emulate pictured objects such as icons on an iPhone, but designers are still able to configure iPhones intentionally, such that represented objects (that are not physically part of the configured surfaces) are emulated. Following the discussion of artefact evolution in Chapter 1, human emulation capabilities (shaped by forgotten conditions of an evolutionary past) now shape the natural selection of graphic representations: following the Playfair example from Chapter 1, graphic representations that use evolved human capabilities effectively are more likely to ‘survive’ than less-effective alternatives.

Returning to pictured objects, we can extend the Chapter 3 discussion of pictured environments to pictured objects in a pictured environment by realizing that authors can use an audience’s naturally-selected capabilities to emulate objects (targets for homeostatic reactions) via pictorial information. Through intentional configuration of an environment (e.g., marks on a surface) to produce pictorial information, an audience’s naturally-selected capabilities to emulate objects (reaction targets) of actual environments can be used to enable emulation of objects that are not part of the configured surface. These are ‘pictured objects.’

The next section will show how graphic representations extend the homeostatic perception-reaction loop by expanding the space for possible reactions, simulations, and memory traces that serve as the resources for possible reactions. This process is important for archives, interfaces, books, etc.

4.3.3 Explanation # 2: Pictured Objects: Homeostatic Reaction Targets

The homeostatic model developed in Chapter 3 can be extended by clarifying how graphically-represented objects amplify/extend the homeostatic process. For example, graphically-represented objects enable a wider range of emulations, thereby enabling more, and more diverse, possible memory traces. Because memory traces are conceptualized as resources for constructing simulations, a greater range of simulations may also be afforded. Finally, because simulations are the source of reactions, a greater range of possible reactions may be afforded.
4.3.3.1 Objects as Targets for Homeostatic Reactions

…the brain interprets relatively abrupt discontinuities – such as change of orientation of a line, change of colour, change of brightness – together as constructing holistic visual objects which are expected to share a ‘common fate’…. No perceptible physical object is ever the same from one moment of its existence to the next. Everything changes. Objects are merely slow events… (Hurford, 2003, p. 276, emphasis added)

As a point of entry to a discussion about objects, I begin by considering Hurford’s statement that “the brain interprets…discontinuities…together as constructing holistic visual objects which…share a common fate…” (2003, p. 276). I will synthesize this explanation with the homeostatic model introduced in Chapter 3 to conceptualize objects as targets for homeostatic reactions.

According to the homeostatic model, an organism’s perception-reaction capabilities were naturally selected to maintain the conditions necessary for the organism to survive (and thrive).26 From an evolutionary perspective, only environmental changes that might require possible homeostatic reactions would be perceptually processed by inherited capabilities (see Section 4.3). Within this context, Hurford’s perceptually-processed holistic objects can be interpreted as the targets of homeostatic reactions.

Example: Natural environment. An orange sits on a table in front of a hungry individual, who reaches for the orange in order to eat it. The orange is perceptually processed as an object because it is a target for a homeostatic reaction (e.g., a plan to eat, move, or manipulate the orange).27 However, suppose the individual instead chooses to pack the orange in a lunch bag for later consumption. Throughout the day, the individual may imagine (simulate) eating the

26 Homeostatic reactions could include ‘internal’ reactions (e.g., releasing a hormone such as adrenalin), reactions that move the organism away from inhospitable locations (e.g., to avoid ‘external’ threats), or reactions that modify the surrounding environment (e.g., building a fire for warmth).

27 Here, a ‘plan’ is a simulated reaction (to an emulated or simulated environment), e.g., a mental simulation that prepares a person to reach for an orange.
orange. In this way, an object can also be simulated (as a simulated target for a simulated reaction). The next section applies this conceptualization of an object to graphic representation.

4.3.3.2 Pictured Objects as Targets for Homeostatic Reactions

An example can help illustrate how pictured objects can be conceptualized as an extension of the homeostatic process: a telerobot controlled by a human operator through an interface. This example will reveal the explicit perception-reaction loop between a remote environment and an operator. It can be extended into domains where the perception-reaction loop is less explicit, but can still explain the behaviour: this is important for archives, repositories, libraries, public policy documents, education, journalistic articles, etc.

**Telerobot.** Consider a radioactive environment, such as Chernobyl, Three Mile Island, or Fukushima. At a time in the not-too-distant past, a human operator wearing protective gear would have been sent into this environment for exploration and cleanup. The (dorsal-ventral) homeostatic process described above would enable this person to perceive-react in this environment, but because the risk is so high, human operators are now replaced by robots whenever possible.

**Single user, single robot.** Consider a simplified model of a telerobotic control loop: A human operator sits at a command and control centre. Graphic displays, composed of graphic representations, enable the operator to perceptually process information about the remote environment and objects in the remote environment. For example, television cameras mounted on the robot might detect reflected light waves in the remote environment and convert these into electromagnetic waves for transmission through cables to the operator’s console. The electromagnetic waves are electromechanically processed in such a way that ‘pixels’ are configured on an information display, such as a television or computer screen. The operator is able to emulate this display information as pictorial information. Based on these emulations of the remote environment, the operator is able to react, and these reactions are transmitted through controls, such as joysticks, buttons, or perhaps haptic devices, via electromagnetic waves, to the robot’s motors and servos.28

28 Note that by my definitions, the ‘objects’ found/constructed in the environment will potentially be different under a telerobotic scenario than they would be if a human were present (because affordances are different, as the
This example illustrates the homeostatic perception-reaction described in Chapter 3, but it does not reflect how telerobotic exploration actually occurs. In real-world scenarios, there are many more ‘users’ of the robot than actual robots, which considerably alters the architecture of the perception-reaction loop, even though the core homeostatic properties remain. The next example will describe this altered P-R loop, which can serve as a template for explaining graphic representation use in other contexts.

**Many users, single robot.** Consider NASA’s recent Mars Exploration Rover mission. Here, an entire scientific community is perceptually processing and reacting to a remote environment via the Spirit and Opportunity rovers (Arvidson et al., 2010). These rovers do have operators who roughly correspond to the operators in the above example, but the true ‘drivers’ of the mission are multidisciplinary science teams in an adjacent operations room. Six or more teams might be in control of various science packages (e.g., sensors and other measurement devices) on the robot. Through a process of negotiation, and a semi-democratic process, (re)action plans for each day are selected and communicated to the rover operators, often through specialized interfaces. Based on these directives, data are retrieved daily, placed in a repository, and displayed through a multitude of devices, thereby altering the perception-reaction loop. Each robot can only execute a finite number of (re)actions in linear space-time; each reaction is informed by perceptually processed information, which serves as the basis for memory traces (resources for simulated reaction plans).

This perception-reaction loop extends even further into society: each science team involved in the mission represents entire branches of scientific communities, such as astrobiology, planetary geology, astronomy, etc. Pictorial information about the remote environment migrates into scientific reports, publications, news articles, and even congressional legislation. A wider multi-national semi-democratic process is involved in selecting larger-scale repertoire of available actions is different). Furthermore, there are different ways to conceptualize the “organism” and thus the notion of “homeostasis”. I will not delve into the complexities here, except to note that we would have to be careful to distinguish among the perception/action of the robot, the perception/action of any particular human in the control room, and the perception/action of the “human-robot-system as a whole”.

---

*This is an automatically generated text. It may contain errors or inaccuracies.*
reaction plans, such as landing sites for current and future robotic probes (see e.g., Arvidson et al., 2008).

This ‘homeostatic loop’ model can also be applied more widely. Consider the famous mutiny on the HMS Bounty (Alexander, 2004); this is a good example because it is readily accessible via popular culture, such as books and movies. The HMS Bounty was able to navigate through waters that had only been explored by Europeans a few times before. This was possible because previous explorers had created detailed graphic representations (records and logs) that served as resources for navigational charts and maps that were replicated and disseminated throughout the British Royal Navy. This example clearly shows how reaction plans (sailing decisions) can be afforded by graphic representations (nautical charts, maps, and instruments). As in the telerobotic example, records, charts, and maps can increase the space for possible emulations (by reading charts), thereby increasing the space for possible memory traces that can serve as resources for an increased space for possible simulations (navigational plans). The increased space for, and accuracy of, possible simulations (navigational plans) increase the space for possible reactions (navigational decisions and actions, such as sailing a ship to barely explored destinations).

After the mutiny, Captain Bligh and loyal crewmembers were set adrift in a small sailing vessel, without the Bounty’s nautical charts and maps. Because Bligh had been to these waters in a previous expedition, he was able to navigate to safety using only his memory, a quadrant, and a pocket watch. Another captain without Bligh’s memory, charts, or maps, would likely have perished. In other words, without external graphic representations, the homeostatic loop (the space for possible emulations, simulations, and reactions) would have been too small to afford the reactions required for survival.

The mutinous crew of the Bounty, led by Fletcher Christian, had less sailing experience than Bligh, and most had never navigated these waters before. However, because they had the Bounty’s nautical charts and maps, they were able to navigate successfully. Interestingly, they happened upon Pitcairn Island, which had been charted incorrectly. Christian realized that this error would have been replicated on all nautical charts and maps throughout the British Navy, and that he and his supporters could successfully hide on this island. He increased his chances of evading the British Naval vessels that he predicted (simulated) would be tasked with finding
them (which would have led to court-martial and probable execution). In other words, Christian surmised (simulated) that his pursuers would use imprecise charts/maps (graphic representations) to plan searches (simulations) that were incorrect, thereby evading capture (and ensuring survival).

This phenomenon appears frequently: individuals and societies that use graphic representations to expand their homeostatic loops survive or prosper more effectively than their competitors (cf. Innis, 1950). For example, a policymaker, scientist, or engineer can more effectively predict the outcome of a decision by using graphic representations that recount past efforts.

The main point here is that pictured objects can be understood as co-opting capabilities for emulating ‘objects’ that are not part of the configured surface (Explanation #1), but the homeostatic explanation (Example #2) can explain, and predict, graphic representation affordances for design decisions across a wider range of cases.

The next section focuses on how pictured objects afford the selection of particular reactions from simulated possible reactions, thus affording this homeostatic reaction loop in both the natural world and designed systems.

4.4 Pictured Attributes

Hurford described the processes that enable perceptual processing of object predicates (attributes), typically associated with the ventral stream, as follows: “[this] process is the somewhat slower analysis of the delivered referent object by the perceptual (visual or auditory) recognition subsystems in terms of its properties” (2003, p. 273).

If an object is a possible reaction target, what perceptually processed information could enable the selection of an appropriate homeostatic reaction (from a range of simulated possible homeostatic reactions)? For example, what information could enable the appropriate grip and pressure for grasping the cylindrical shape and heft of a hammer handle (see Figure 4-4, left), relative to the spherical shape and lighter weight of an apple (see Figure 4-4, right)? Perceptually processed shapes, colours, and textures afford the recognition required to configure and select appropriate homeostatic reactions. For example, an individual might visually process an apple’s mouldy texture and colour during object recognition. The recognition process uses ventrally
connected memory traces of prior perceptions, which are used as the resources for the construction of simulations or re-enactments, e.g., the bad taste of a rotten apple eaten as a child. This recollection may make the individual decide to throw the apple into the garbage, rather than eat it.

Thus, the organism’s homeostatic perception-reaction to attributes (e.g., shape, texture, colour) unifies the (network of) attributes into a whole object. Perceptually processed object attributes and relations serve the functional purpose of enabling an appropriate reaction.

The next sections will explore how authors of graphic representations can use these capabilities to picture attributes of objects (such as shapes) or details about objects (such as textures, colours, or luminance).

![Diagram of a hand holding a hammer and an orange]

Figure 4-4: Attributes as information that enable appropriate reaction to change-variation.

4.5 Symbolic Objects (Afforded by Pictured Attributes)

Attributes of an object in a natural environment enable a perceptual system to recognize, and therefore distinguish, objects in an environment. Objects are reaction targets, so distinguishing among objects is a way to distinguish among possible (simulated) reactions to targets. In this way, a natural object’s attributes afford the construction of simulated possible reactions from memory traces that are activated while recognizing the object. This section explores how pictured attributes extend this process into the domain of external graphic representation.
Figure 4-5 is a screen capture from Microsoft Word. Each button-icon, a graphic object, is a target for a reaction: clicking (reacting to) one button-icon (e.g., “Undo”) instead of another (e.g., “Tables”) will produce different outcomes. How does one’s perceptual system distinguish among graphic objects in order to simulate and select an appropriate reaction (button click)? Distinguishable properties, such as shapes and colours, enable the user’s perceptual system to differentiate among these graphic objects, and (via ventrally connected memory structures) to simulate the possible outcome of reacting to a particular icon by clicking it. Let us now consider what can be simulated when a graphic object is recognized.

Figure 4-5: Icon-buttons in Microsoft Word: Graphic Objects?

Section 3.1 described how a simulation can be a re-enactment of a memory trace of a prior percept, or an ‘abstraction’ of memory traces that emerge via repeated re-enactments. This means that objects and abstracted objects can be simulated, as can object attributes and abstracted attributes, and relations and abstracted relations among objects. Attributes of a pictured object can be intentionally configured by an author to afford, through recognition, intended simulations of objects, attributes, or relations. These are symbolic pictorial objects.29 The next section will synthesize the theory of symbolic information presented in Chapter 3 with the current conceptualization of object attributes and how these can describe symbolic pictorial objects. Hereafter, symbolic pictorial objects are referred to as ‘symbolic objects’ or ‘symbols.’

4.5.1 Example: Text-Sentences Part A

Let us now explore some examples of symbolic objects affording recognition (and the construction of simulations from memory traces of past percepts), using the conceptualization of ‘representation’ introduced in Chapter 3 (see Section 3.3). Consider the printed word

29 Here I must distinguish between the graphic representation and that which is represented. The symbolic graphic object can represent simulated objects, simulated relations, or simulated details. The attributes of the symbolic graphic object’s pictorial properties are what afford recognition, and therefore, the construction of intended simulations.
“recognition.” It appears as a row of symbolic objects (letters/graphemes) in a linear arrangement: “r-e-c-o-g-n-i-t-i-o-n.” Each letter has some common attributes: e.g., size and colour, which might afford their recognition as a subclass within the ‘universe’ of different kinds of graphic representations: the text-alphabet (as opposed, for example, to electrical symbols or architectural notations). Additionally, the shape attributes of each letter-grapheme are such that my perceptual system processes them as a row of ‘distinct’ graphic objects: individual letter-graphemes of the text-alphabet.30 Let us now consider the hierarchical nature of this recognition process that appears to be afforded by these common and distinct shape attributes, using the metaphor of an address system. I use the example of a zip code: 15217, the zip code of Carnegie Mellon University (CMU), in Pittsburgh, Pennsylvania, USA. In zip code 15217, the first character (“1”) specifies a broad region of the United States (Delaware-New York-Pennsylvania). The next two characters (“52”) specify a city (Pittsburgh). The final characters (“17”) presumably specify a neighbourhood (e.g., Oakland, where CMU is located). Similarly, the common shape attributes of all letter-graphemes of the text-alphabet seem to afford their recognition as members (letter-graphemes) of the text-alphabet, a subset within the universe of other kinds of symbols (electrical symbols, architectural symbols, etc.), whereas the individual shape attributes of each letter-grapheme (e.g., the ink outlines on a page that represent the letter “r”) affords (via the ventrally connected recognition capabilities just described) the identification of particular letter-graphemes among the set of all letter-graphemes of the text-alphabet.

Consider how this hierarchical, and somewhat standardized, recognition capability must have been intentionally constructed. More specifically, consider how this recognition capability must have been constructed, such that recognition capabilities across individuals had enough common properties to enable an author’s intentions to be simulated by audiences when they recognize a graphic object that has been intentionally configured by an author.

Chapter 3’s conceptualization of external graphic representation provided the foundation for considering how this recognition capability might have been constructed, by building on the conceptualization of recognition presented in Chapter 2.

30 Note how printed letters on a page seem ergonomically configured to fit the highest frequency zone of the retina/visual cortex, and the region most associated with recognition.
A formalized, cultural learning process, primarily facilitated via standardized educational materials, causes the intentional construction of memory traces among learners. Furthermore, structured curricula with approximately common properties across institutions cause reenactments of these memory traces across learners, resulting in hierarchical ‘memory trace abstractions’ with common properties among learners. As a result, it is possible for me, the author of this text, to simulate (with a level of certainty higher than chance) how you, the reader, might recognize the graphic object properties (letter-graphemes) that I have used on this page. Likewise, it is possible for you, the reader, when recognizing the shapes of the individual graphic objects that I have used, to simulate (with a level of certainty higher than chance) what I intended, when I intentionally used these graphic objects on this page.

At this point, it is almost time to explore how pictured and symbolic objects can be arranged in configurations to produce pictures, diagrams, and sentences, by pictorially representing (or ‘picturing’) objects and relations. However, before discussing hybrid graphic representations that are composed of pictorial or symbolic relations among objects, it is important to clarify what I mean by ‘relation.’

4.6 Pictured Relations

The previous few sections were conveniently able to adapt Hurford’s account to objects and attributes. Unfortunately, Hurford did not appear to extend his discussion to relations. The quest to conceptualize relations spans the history of cognitive science and philosophy and cannot be fully treated here (see Jung & Hummel, 2013 for a recent example). However, a brief characterization within the terms introduced thus far will help us conceptualize affordances of pictorially and symbolically represented relations in the next subsections. I will conceptualize relations in this chapter as the information that enables perception of spatial-temporal configurations among differentiated target objects. The above conceptualization is elucidated in the following manner:

Let us begin with everyday intuitions about objects. Intuitively, an object cannot be in more than one place at one time (more than one place in space-time). Let us now consider the homeostatic conceptualization of objects as targets for perception-reactions: Reactions directed at a particular target transpire at given subsets of space-time.
Within a given subset of (or point in) space-time, an Object A, in an environment shared with Object B and C, is in relation to Objects B and C. Whereas objects are targets for reactions, and attributes (such as shape, colour, or texture) inform the simulation and enactment of reaction plans to target objects, relations inform perception about the configurations among the differentiated target objects. Within this simplified conceptualization, the very capability to perceptually process an object requires a capability to distinguish it from, and in configurations with, other objects.

Let us further develop the above by considering evidence from the field of developmental psychology that reveals how babies can perceive relations among objects only after they are able to perceive objects (see the work of Jean Piaget for classic examples). The following thought experiment applies the simplified account of recognition outlined in Chapter 2 (based on Barsalou, 2009). Imagine a simplified organism that is trained through example directly after it is ‘born.’ We show it an object (Object #1) for the first time. Given that the organism was just born, this object comprises the majority of its life experience. Memory traces of Object #1 are constructed. Next, we show it another object (Object #2); the organism’s perceptual processing of Object #2 uses the memory traces left during the perceptual processing of Object #1, and also leaves additional traces. So, some traces are in common between Object #1 and #2, and some traces are not in common between Object #1 and #2. The traces that are not in common are the basis for the organism’s emerging capability to distinguish/differentiate among objects (e.g., Object #1 from Object #2). These emerging capabilities help the organism to distinguish relations among objects in space-time (for a more complete discussion, see Barsalou, 2005, p. 400-408).

Next, consider an apple and an orange in a natural/actual environment. My perceptual system enables me to process relations in several ways. A) I can distinguish among the spatial relationships of the apple relative to the orange. B) I can distinguish relationships among the colours and shapes of the respective objects. C) Conceptual relations are afforded, if e.g., the apple conjures a happy memory, and the orange conjures a negative memory. A-B can be emulated from the natural environment, whereas C must be simulated. The next section provides a brief discussion of emulated relations (A-B) as the basis for pictorial information. The following section will discuss C (simulated relations), specifically with regard to how pictured
objects can be intentionally configured to cause intended simulated relations (as well as simulated objects and their attributes): these are symbolic graphic objects.

**Example of Pictured Relations.** Figure 4-8 is a pictorial graphic representation that enables a bottom-up perceptual processing (emulation); in turn, this affords the perceptual processing of *pictured relations* among *pictured objects*. Authors of graphic representations can make use of this capability, as the author Figure 4-8 did, enabling emulation of pictured spatial, colour, and shape relations.31 Pictured relations of shape, colour, and space can also be used symbolically (see Figure 4-10). Let us now expand on item C, above, to discuss simulated relations, the basis for symbolic relations.

31 To further demonstrate the notion of pictorially represented relations, in Figure 4-6, the dark red areas of Frame B highlight the pictured relations of Frame A, whereas the dark red areas of Frame C highlight the pictured objects that the pictured relations are among. Of course, if the water of Frame A (highlighted in Frame B) were the target object of attention, then the highlighted areas of Frame C could be taken as pictured relations among the highlighted water of Frame B.
Figure 4-6: Demonstrating pictorially represented relations among pictorially represented objects.

Figure 4-7: Relations in information visualization.
4.7 Symbolic Attributes and Relations

Symbolic Relations. As mentioned in Sections 4.5 (and extended in Section 4.6), the attributes of a graphic object can cause simulations of relations to be constructed in the mind of a perceiver. These can also be abstract relations (vs. concrete relations), such as left, right, up, down, beyond, over, etc. (why these are treated as abstract here is detailed in more depth in the next chapter)

Symbolic Attributes. Additionally, as mentioned in Section 4.5, the attributes of a graphic object can cause simulations of attributes other than what is emulated from the graphic representation’s configured surface to be constructed in the mind of a perceiver. These can also be abstract (vs. concrete attributes) attributes, such as red, blue, green, round, square, etc.

The next section summarizes this conceptualization of objects, attributes, and relations, and applies it to describe how pictorially representing (via emulation) or symbolically representing (via simulation) object relations, shapes, and details can describe graphic representation types that afford different kinds of perceptual-cognitive reactions. These are realistic pictures, outline drawings, diagrams, and sentences. This way of describing relations, shapes, and details as pictured or symbolically represented in order to distinguish graphic representation types is a central contribution of this dissertation, and sets the stage for predicting perceptual-cognitive affordances of graphic representation types.

4.8 Synthesis: Hybrid Graphic Spaces

This subsection applies the conceptions of object, attributes, and relations described in the sections of this chapter leading up to this point in order to provide an account of graphic objects, graphic relations, and graphically represented attributes.

4.8.1 Graphic Objects versus Graphic Spaces

Graphic object. The conceptualization of graphic objects builds upon the notion of pictorial information developed in Chapter 3. Here, a graphic object is conceived of as a pictured object. In other words, a graphic object is perceptually processed when pictorial information (from a graphic representation) engages an organism’s capability to perceptually emulate an object within an environment (e.g., perceptually emulating change-variation to simulate a
possible reaction to that change-variation). For the purposes of this discussion, a pictured apple that is processed as an object is an example of a graphic object.

Pictorially representing an object means that attributes must be pictured graphically (e.g., shape, colour, and texture). Pictorially representing relations is possible when a graphic representation pictures graphic objects in relation to other graphic objects – these are graphic relations.

**Symbolic graphic object.** The following conceptualization of a *symbolic* graphic object builds upon the conceptualization of symbolic information presented in Chapter 3. Here, a symbolic graphic object is a graphic object that, via learning, causes retrieval of simulated non-pictured environment. For example, the logo of Apple Inc. is a pictured apple, but through learning, a viewer can recognize the attributes of the pictured apple as referring to Apple Inc., which is not pictured. A perceiver cannot perceptually emulate Apple Inc., the *abstract* concept, because it is more than a physical location with buildings. Let us say that the pictured apple symbolizes Apple Inc. Thus, a particular graphic object can symbolically represent relations and object attributes that are not pictured. From this perspective, pictorial (or graphic) information can cause the emulation of objects, relations, and attributes. Additionally, symbolic pictorial (or graphic) information can cause the construction of simulated objects, relations, and attributes. Framing graphic representations in this way enables the conceptualization of hybrid graphic representations, such as:

- Pictured relations among pictured objects as realistic pictures or outline drawings (Figure 4-8; Table 1, Column 2);
- Pictured relations among symbolic objects as diagrams (Figure 4-9; Table 1, Column 3);
- Symbolic relations among pictured objects as, e.g., comic strips, certain kinds of flow charts (Figure 4-10; and
- Symbolic relations among symbolic objects as sentences (Figure 4-11 and Table 1, Column 4).

---

32 ‘Pictorially representing’ (or ‘picturing’) is not to be confused with ‘picturing in your mind’. This is specifically referring to pictorial information of external graphic representations that make use of emulation capabilities.

33 Previously I defined ‘graphic object’ in terms of emulation. A symbolic graphic object involves (by its nature) both simulation (for symbolic information) and emulation (for pictorial information). From this perspective, all graphic objects are composed of pictorial and symbolic information, but to varying degrees.
**Graphic spaces.** These hybrid graphic representations will be described as *graphic spaces*, an extension of Engelhardt’s use of the term. Specifically, a ‘graphic space’ is a representation composed of graphic objects, relations, and attributes, whether they are pictured or symbolic. Examples could include a sentence, a figure in a magazine, or an academic paper.

![Figure 4-8: A realistic picture as pictured relations among pictured objects.](image1)

![Figure 4-9: A diagram as pictured relations among symbolic objects.](image2)
Diagrammatic graphic spaces. As introduced above, a graphic representation can picture relations among symbolic graphic objects (Table 1, Column 3). Here, I refer to these pictured relations as diagrammatic information, and a graphic space composed of diagrammatic information as a diagrammatic graphic space. In this type of graphic representation, relations are emulated from pictured relations among symbolic graphic objects. For example, consider a stock-market chart that depicts pictured relations among symbolic objects. A chart could show the value of Apple Inc. relative to Microsoft. Apple Inc. and Microsoft are presented through
symbolic graphic objects (words or logos), but the relations between them are pictured. Consider this litmus test: if the font of “Apple” were changed, the meaning would not change. However, if the pictured relation changed, the meaning would change.

**Sentential graphic spaces.** As introduced above, a graphic representation can symbolize relations among symbolic graphic objects (Table 1, Column 4). Here, the symbolic relations among symbolic graphic objects are referred to as sentential information, and a graphic space composed of sentential information is referred to as a sentential graphic space. For example, consider the following sentence: ‘The apple is to the right of the orange, and both are on the table.” Here, “apple” and “orange” are symbolic graphic objects, while “to the right” and “on the” are symbolic relations.

**Recursion.** Graphic spaces can be considered graphic objects, and thus composed into other graphic spaces. Graphic objects can be decomposed into graphic spaces and graphic spaces can be taken holistically, as graphic objects.

### 4.9 Summary and Conclusions

Chapter 4 introduced the idea of distinguishing graphic representations in terms of how graphically represented objects and relations can be pictured or symbolized. How is this significant? The idea builds substantially on the conceptualization of pictorial and symbolic information presented in Chapter 3, and makes it possible to define hybrid graphic representations, such as diagrams, as pictured relations among symbolic objects. Additionally, the idea enables sentences to be conceptualized as symbolic relations among symbolic objects, and pictures to be conceptualized as pictured relations among pictured objects. To achieve this configural approach to distinguishing graphic objects, this chapter introduced the idea of a graphic object as an intentionally configured representational object. From this perspective, representational objects increase human survival and prosperity by extending their homeostatic perception-reaction loops in such a way that a greater variety of reactions to a greater variety of possible environmental change-variations is afforded. The significance of this explanation is that it enables an intersection between ideas about human natural selection and artefact natural selection. For example, graphic representation properties and types are naturally selected because they afford human survival and prosperity. When configurations of humans and their artefacts
enable survival and prosperity more effectively (e.g., requiring less energy; see e.g., Odum & Hall, 1995) than competing configurations, they are more likely to survive and thrive.

**First Half**

Let us now discuss the significance of this capability to distinguish and describe graphic representation properties in relation to the overall goals of the first half of this dissertation. Chapter 1 discussed how visual information design is currently practiced as a craft; now a scientific approach (a model) is needed. Chapter 1 extended this discussion by exploring how a scientific model enables predictive explanations. This required ways of describing phenomena, which were identified as Type I or Type II theories. Thus, the first half initiated the process of establishing a way to describe graphic representation phenomena, as a first step in a journey that could lead to predictive scientific explanations.

Chapter 1 also introduced a methodological approach (or guiding theme) called ‘sites of artefact evolution’ (SAE). The goal was to identify proto-theories that emerged from practice, and to apply these as a way to identify concepts and key terms from the cognitive neurosciences. This helped establish a vernacular to act as an arbiter between the applied arts and the cognitive neurosciences.

Chapter 3 was the culmination of this effort to establish the key terms, and established the concepts these referred to in order to describe the properties of graphic representations typically used in visual information design, but using perceptual-cognitive terms.

**Status Relative to Theory Goal**

This concluding section marks a transition between two kinds of projects. Now that a language for describing graphic representation properties has been initiated, the objective is to demonstrate that language by describing graphic representation phenomena and behaviours [affordances]. The next half is more speculative than the first, and the bulk of my career will likely be spent refining what follows. The main objective of the second half is to demonstrate how the model can be used to explain phenomena, and how those explanations can be correct as well as incorrect.
The next chapters will take a small step toward demonstrating the model, by providing an explanation for the properties and affordances of graphic representation types that were naturally selected in Euclid’s Elements. This is a behavioral demonstration that can be explained using the vernacular developed here to discuss what is going on ‘under the hood.’ In turn, this will set the stage for tentative predcations about the properties and affordances of graphic representation types more generally.

**Growing Theme: Culture**

One issue lurking in the background throughout the first half, and especially Chapters 3 and 4, is the role played by culture, and cultural institutions, in shaping the shared intentions used by graphic representations. This issue will become clearer as I describe the natural selection of sentences from the ‘primordial soup’ of pictures, and how those sentences afforded the communication of abstract concepts that cannot be pictured, thereby setting the stage for the emergence of the sentential axiomatic method. Sentences, and the axiomatic method they support, afford the communication of abstract concepts that are difficult to picture, and they are the basis for the predictive power of modern science, which dramatically extends the homeostatic reaction loop.
5 Predicted Affordances of Graphics

5.1 Predicted Affordances of Pictorial and Symbolic Information

5.1.1 Introduction

This chapter serves two purposes. Its primary purpose is to show how the affordances of pictorial and symbolic properties and pictorial and symbolic information can be predicted. Its second purpose is to integrate information theory into the previously introduced concepts of pictorial versus symbolic. A brief review of theories about perceptual and conceptual categorization will demonstrate how the development of categories enables the properties of graphics or an environment to be processed with greater certainty or specificity (providing a bridge to information theory). This review of perceptual and conceptual categories will also introduce a working approach for identifying the concepts or structures that an author intended to graphically represent as ‘concrete’ or ‘abstract,’ based on how modally specific (e.g., visual) or amodal the intended concept or structure is. By integrating these concepts (of information theory and categorization) into the model, it is possible to predict how pictorial information affords the sharing of concrete structures (or concepts) and how symbolic information affords the sharing of abstract concepts.

5.1.2 Pictorial and Symbolic Information: Perceptual and Conceptual Categories

Previously, I discussed how the model can conceptualize perceptually processed information (of a graphic or environment) as a function of an organism’s perceptual capabilities. This subsection will explore pictorial and symbolic information in more detail, by delving deeper into the structure of the human perception-reaction system, and integrating theories of so-called ‘perceptual’ and ‘conceptual’ categories (Mandler, 2006) into the perception-reaction model that serves as the foundation for the model. To this end, I will expand on a theory referred to as ‘intellego perception,’ set out by Norwich (1991), according to which information is a function of perception. Norwich’s theory expands on Shannon and Weaver’s (1948) work by applying it to perception. The idea of ‘certainty’ is central to this discussion, so let us begin by clarifying what is meant by ‘more or less certain’ in this context:
... entropy, taken as a measure of the absence of knowledge (following Shannon 1948),
can have a well-defined meaning only in relation to a space that...is properly
“regularized” or that (in the words of Wiener 1948) has a “fundamental equipartition.” To
take the simplest example, if we have no knowledge about the location of a point in a
one-dimensional space, we can only suppose that every location on the line is equally
probable. (This is the “principle of indifference” so successfully employed in physics by
Maxwell and Boltzmann – see Jaynes 1978.) (Shepard, 2001, p. 600)

Consider the line shown in Figure 5-1 (right). Here, we have no knowledge about the
location of a point in a one-dimensional space, and every location on the line is equally probable.
Compare this to the line shown in Figure 5-1 (left). Here, we have some knowledge about the
location of a point in a one-dimensional space, due to the shaded red marker. This means we
have more certainty (or more information) about the location of the point in Figure 5-1 (left) than
we do about the location of the point in Figure 5-1 (right).

The idea of certainty can also be applied to the human perceptual system. Norwich (1991)
extended Shannon and Weaver's (1948) account of information to explain how (sensory)
information is not only a measure of uncertainty, but also a function of perception:

...the precision, and even the existence, of a percept is dependent on the richness of the
set of possibilities that was available to the perceiver. That is, if the percept had not been
a part of the set of alternatives it could not have been made, and if the set had consisted of
very few members the percept would, in some respects, be less unique than if the selection had been made from among many members…. …The set of alternatives, in turn, is dependent upon the physiological structure and history of the perceiver. The perceiver…who chose “blue” could not have done so if he had lacked “blue” among his experienced colours or lacked blue photoreceptors. We can capture this idea by describing the perceptual process as relative. A percept is relative to the physiological function and history of the perceiver and, of course, requires the existence of a memory (Norwich, 1991, p. 82).

Basically, an action or item is selected or identified relative to a range of actions or items that are not identified. For example, suppose I am on a desert road and see a speck on the horizon (Figure 5-2a). The speck is aligned with the road, and is growing larger (Figure 5-2b). I realize that a vehicle is on the road and moving in my direction.

My prior experience in the world and with roads has taught me that the vehicle could be classified as one of a range of possible categories, e.g., ‘motorcycle,’ ‘sedan,’ ‘truck’ (or ‘van’), or ‘18-wheeler.’ Additionally, my experience has taught me that the vehicle is unlikely to be classified as a number of other conceptual categories: e.g., ‘house,’ ‘person on foot,’ ‘aircraft,’ or ‘Sherman tank.’

![Figure 5-2: Uncertainty and certainty applied to perception.](image)

I wonder: "Is it a motorcycle, sedan, truck, or 18-wheeler?" As the vehicle draws nearer and additional features become visible, I conclude: "It is either a car or truck, but not a motorcycle or 18-wheeler" (Figure 5-2c). The object grows closer still, and I see that it is a sedan, but not a motorcycle, truck, or 18-wheeler (Figure 5-2d). The conceptual category of the incoming object is identified with more certainty as it draws closer.
However, to relate information to both pictorial and symbolic properties of a graphic (and relative to both emulation and simulation), one dimension (e.g., more or less certainty) is not sufficient. Recall the Chinatown example: even though I might not know what conceptual category a written Chinese symbol falls under, I can still process the pictorial properties (e.g., Gibson’s optic array) of the marked surface. Additionally, consider the example of a medieval chapel: an audience not only understands the religious significance (the author’s intended ‘symbolic meaning’) of, say, a painted sheep, but also sees the brushstrokes of the marked (pictorial) surfaces. The next subsections present two dimensions to describe information that is perceptually processed via the (emulated) pictorial and (simulated) symbolic properties of a graphic.

5.1.3 Perceptual versus Conceptual Categories

This section discusses how perceptual and conceptual categories can be introduced into the model of perception-reaction described above.

Mandler described two kinds of categorization capabilities that develop in infants: “Perceptual categorization” is used to compute the perceptual similarity of objects; it may be the foundation upon which infants “attempt to make sense of what they perceive – to construe the meaning or significance of the events they observe” (Mandler, 2006, p. 3). These construals are thought to lead to more explicit ‘rule-based’ conceptual categorization. Let us now consider the physical structure of these processes in relation to the model of perception-reaction.

Recall that each perception-reaction leaves memory traces in the form of conjunctive neurons across lower-level association areas (the left side of Figure 5-3 and 5-4). These memory traces become resources for simulating (or ‘prototyping’) potential perception-actions, by re-enacting prior perception-reactions. At lower-level association areas that are more tightly coupled with sensory receptors, simulated prototypes fall under perceptual categories, using many of the same perceptual systems used for the original percept that serves as the basis for later re-enactment.
At higher-level association areas (the right side of Figure 5-3 and 5-4), conjunctive neurons converge in zones across multiple sensory modes. These ‘convergence zones’ (Damasio, 1989; Simmons & Barsalou, 2003) enable simulated prototypes of possible perception-reactions that are not as easily described in terms of a specific perceptual mode or a reenactment of a specific prior perception-action. Instead, these simulated prototypes fall under more general categories of possible perception-actions (Barsalou, 2003). These are not only more amodal, but have been described as more filtered, interpreted (Pylyshyn, 1973), conceptual (Barsalou, 2003; 2005), or abstract (Barsalou, 2003). For example, a young child who takes a bite out of what
turns out to be a rotten apple might latter re-enact those experiences when she perceives another rotten apple with common properties. Over time, she develops an understanding of ‘rotten’ as a category that can include apples, as well as many other objects, people, and experiences. Similarly, a child can learn to associate sounds with certain intended meanings (learning a language), or marks with intended meaning (learning to read). The abstract concept of ‘square’ can apply to a shape on a raised surface that is touched but not seen, as well as to a drawing on a piece of paper that is seen and not touched. These less-modally-specific simulations have been described as more ‘interpreted’ or ‘conceptual,’ and more perceptually-based simulations are considered to be more ‘concrete.’

I will now focus on how graphics use perceptual and conceptual categories. Specifically, I will refine the two technical terms used in the model to denote aspects of this perception-reaction loop so that pictorial and symbolic properties of graphics can be distinguished and described.

Returning to external graphic representations to employ the concept of information introduced above, consider the perceptual processing involved in the Chinatown example, where the writing was customized for audiences based on cultural-linguistic backgrounds. To perceptually process writing, individuals emulate light structures produced by a source, such as a mark on a surface. Emulated properties from the graphic are processed and filtered through memory structures to simulate a prototype that falls under an intended conceptual category. Each individual's experiences (especially cultural interactions) help them develop conceptual categories that shape how these emulations are filtered and processed to construct simulated prototypes of an author's intended meaning. As categories become denser and more numerous, more nuanced (specific or certain) simulated prototypes can be constructed and communicated. However, because conceptual categories build directly or indirectly on perceptual categories (which are more closely linked with sensory receptors that process information from the physical world), conceptual categories vary more among individuals than perceptual categories do.

Within the context of describing and distinguishing information of an external graphic representation, the two examples reveal two different dimensions. The first dimension is a function of the conceptual categories that build on the lower-level perceptual categories that are used during the perception of a graphic: symbolic information. The second dimension is a
function of the lower-level perceptual categories that are closely linked with the sensory receptors that are used during perception of a graphic: *pictorial information*.

For example, let us consider the case of picture perception in the Chinatown example: individuals who speak, write, and understand different languages are still able to use the pictures. Lower-level perceptual categories may enable some pictures to be used across cultural and linguistic differences, because of two interrelated factors. First, individuals from different cultures inherit perception-reaction capabilities from common ancestors. This means that many biologically grounded capabilities (e.g., the human eye) have common properties across diverse cultures. Second, optical structures produced by reflected light from physical objects, such as occluded surfaces and edges, are emulated and processed using the inherited perception-reaction capabilities, and help construct perceptual categories. As a result of these two interrelated factors, individuals worldwide can perceive and act within physical environments using their developed perceptual categories. When an artist who speaks and writes in a language that is unfamiliar to her audience creates a picture (by placing marks on a surface to produce optic arrays/structures, which use her audience's lower-level perceptual categories that developed through their interactions with physical environments), the audience can use the picture regardless of linguistic background. Each individual's inherited capabilities (with common perceptual-action capabilities) develop perceptual categories that are closely linked with sensory receptors. As the perceptual categories grow denser and more numerous during development, more nuanced emulated and simulated prototypes of lower-level (concrete) features can be emulated or simulated. Simply, the individual is developing the capability to process pictorial properties with greater certainty: the capability to process pictorial information.

As mentioned above, both types of information are a function of perception, and are shaped by inherited biologically grounded capabilities that help an individual develop perceptual and conceptual categories over a lifetime. However, compared with symbolic information, pictorial information relies more on perceptual categories that develop within physical environments. Conversely, compared with pictorial information, symbolic information is more shaped and interpreted by the individual and his/her (often culturally grounded) experiences.

If pictorial properties are emulated properties of a graphic, then pictorial information is the information acquired when perceptually processing pictorial properties of a graphic.
However, by introducing information as a reduction of uncertainty, the model can define 'pictorial' in terms of certainty. ‘More pictorial’ information means that an audience has more certainty about the optical structure that an author intended the audience to emulate (C1: Pictorial Information Claim). However, this phrasing can be unwieldy: an author is not aware she is trying to produce a specified optic array! Instead, she is trying to communicate an item via a picture. To simplify our vernacular, it is helpful to apply Barsalou's notion of ‘conceptual’ as more amodal and abstract, whereas ‘more modal’ or ‘more perceptual’ is more concrete. The next subsection will focus on this strategy.

5.1.4 P1: Concrete Pictorial Prediction:

The first prediction of the model, the P1: Concrete Pictorial Prediction, anticipates that pictorial information affords the communication of concrete concepts (or structures) with more certainty than symbolic information. Recall that emulation occurs at the sensory surface where sensory stimuli/sensory information is picked up (Figure 5-4, left). Conjunctive neurons that are closely coupled to the sensory surface become resources for simulating potential perception-actions. At higher-level convergence zones, simulations are increasingly amodal (Figure 5-4, right). Extrapolating from Barsalou’s (Section 5.1.2) conceptualization, emulation is therefore more concrete than simulation (and more modal simulation is more concrete than more amodal simulation). Because pictorial information is conceptualized as the emulated information of a graphic (Claim 1; e.g., when an author configures a graphic to cause an intended audience emulation), pictorial information is therefore predicted to inherit the concrete (more modally specific-precise) properties of emulation. From this perspective, pictorial information is predicted to afford the communication of concrete structures more effectively (with greater certainty) relative to abstract concepts.

Figure 5-5 helps illustrate the P1: Concrete Pictorial Prediction by extending the line example from Section 5.1.2, where information is conceptualized as a reduction of uncertainty. The left line in Figure 5-5 corresponds to emulation at the sensory surface where sensory stimuli/sensory information is picked up. Each point on the left line refers to a possible emulation of a concrete structure of an environment. When emulations are processed and filtered to produce more amodal simulations, they fall within a range of possible conceptual categories, indicated by the right line in Figure 5-5. Each point on the right line refers to a possible conceptual category that a more amodal simulation could fall within.
Let us now apply the foregoing to pictorial representation: The highlighted segment of the left line indicates an emulated concrete structure that could be intended by an author (by configuring a graphic to produce pictorial information). When emulations are processed and filtered to produce more amodal simulations, they fall within a range of possible conceptual categories, indicated by the highlighted segment of the right line in Figure 5-5.

During perceptual processing, an emulation of a given (more certain) concrete structure from a physical environment (the highlighted segment of the left line, indicating a more certain concrete structure) can engender a wider range of possible more amodal simulations that could fall within a range of possible abstract conceptual categories as it is processed, filtered, and interpreted (the highlighted segment of the right line, indicating a less certain abstract concept). For example, a concrete structure (a more particular tree, left) could fall within a range of potential conceptual categories (right; ‘tree,’ ‘nature,’ or ‘autumn’). Similarly, in a given graphic, when pictorial information conveys a concrete structure with more certainty (left line), the abstract conceptual category within which the pictorially conveyed concrete structure could be processed to fall is less certain/more ambiguous (right line). For example, the pictorially represented more concrete structure (a depiction of a more particular tree, left) could fall within a range of potential conceptual categories (right; ‘tree,’ ‘nature,’ or ‘autumn’). In other words, when pictorial information increases (when a graphic conveys a concrete structure with more certainty), the graphic is more conceptually ambiguous (because the pictorially represented concrete structure can fall within numerous potential conceptual categories). What this means is that if an author is more interested in conveying a particular concrete structure (and less interested in conveying an abstract concept), she should configure an environment (e.g., by placing marks on a surface) to produce light structures that she anticipates will be emulated by an audience (as pictorial information). Because audience emulation is more tightly coupled with the pickup of the light structure produced via the author’s intentionally configured marked surface, the structure is more likely to be processed as the author intended. In other words, it is shaped less by prior experiences and more by the emulated structure of the environment that was configured by the author.
Figure 5-5: In this demonstration of the Concrete Pictorial Prediction (Prediction 2), a pictorially represented more concrete structure (a depiction of a more particular tree, left) symbolically represents a less certain, more abstract range of potential conceptual categories (right).

Recall the Chinatown example: I can emulate pictorial properties of both the photographs of food and written labels. Because pictorial properties use lower-level perceptual categories that develop while perceiving and acting in the real world, the pictures afford my actions when the pictures use the perceptual categories I have developed, which are closely coupled with my inherited visual perception capabilities. Emulated pictorial properties of the written graphics also use lower-level conceptual categories. To illustrate this emulated aspect of the written Chinese graphic, consider how a perceiver who is unfamiliar with the Chinese language could trace the inked writing with their hand. However, they would not have developed the capabilities (shared intentionality, as described in Chapter 3) required to perceive the author’s intended symbolically represented item.

The way that the process of artefact evolution has shaped graphic representation practices (when the functional purpose is to communicate concrete structures) can informally demonstrate the P1: Concrete Pictorial Prediction. For example, when one of the primary functional purposes of a graphic is to communicate the concrete structure of a shoreline to a ship navigator, the pictorial properties of an outline drawing are typically employed (for example, note how the
pictorial properties of Figure 5-6 communicate the concrete structures of a shoreline deemed relevant to ship navigation).

Figure 5-6: The concrete structure of a coastline conveyed via pictorial properties of the outline drawings of a map (a).\(^{34}\) The concrete structure of an ocean floor conveyed via pictorial properties of the shaded regions of a map (b).\(^{35}\)

\(^{34}\) http://www.usna.edu/Users/oceano/pguth/website/shipwrecks/intro_lesson/old_chart.png

\(^{35}\) http://www.opc.ca.gov/webmaster/_media_library/2010/02/nautical-chart-corrections2-1024x648.jpg
However (and as will be noted in the next subsection’s discussion of the model’s second prediction about the affordances of symbolic properties) the *names* of cities, countries, and regions (these are abstract concepts, according to this chapter’s interpretation of Barsalou’s conceptualization) are labelled through symbolic graphics such as text (also see Figure 5-6a). Looking at Figure 5-6a, one might think that the numeric symbols for depth (“2”, “2”, “5”, “4”, etc.) represent a fairly concrete thing, namely the precise depths at those points. However, it should be noted that even if so, there is a great deal of uncertainty about the depth in regions between individual numbers. For example, between the “2” and the “4” in the lower right corner, we might assume that the gradient is gradual, but this could lead to a grave mistake if our ship requires a minimum depth of 3 and the actual situation is that there is a shelf of depth 2 in most of the region between the “2” and the “4”. So if our purpose for the map included conveying precise depth information, this dissertation’s prediction would be that evolution would favour a more pictorial representation of the concrete depth, such as the example shown in Figure 5-6b.

Various fields and subfields specialize in the art and craft of communicating concrete structures through pictorial graphics. In addition to the map-making practices used in cartography, the field of biomedical illustration has flourished through the development of pictorial representation practices; pictorial representations now serve as the centerpieces of many biomedical publications (see Figure 5-7).
Figure 5-7: The art and craft of biomedical illustration specializes in the communication of concrete anatomical structures. This illustration is by Dorcas Hager Padget, who was a staff medical artist in the Medical Art Service Department at the University of Toronto in 1930.36

5.1.5 P2: Abstract Symbolic Prediction

Let us now consider the model’s second prediction, the P2: Abstract Symbolic Prediction. It anticipates how symbolic information affords the communication of abstract concepts with more certainty than pictorial information. Recall how conjunctive neurons that are closely coupled to the sensory surface become resources for simulating potential perception-actions and at higher-level convergence zones, simulations are increasingly amodal (Figure 5-5, right). Extrapolating from Barsalou’s conceptualization of concreteness versus abstraction (Section 5.1.2), simulation is therefore more abstract (relative to emulation). Because symbolic information is conceptualized as using simulation capabilities (e.g., when an author configures a graphic to generate an intended audience simulation), symbolic information is therefore predicted to inherit the abstract (less modally specific-certain) properties of simulation. From this

perspective, symbolic information is predicted to afford the communication of abstract concepts (with more certainty relative to pictorial information).

Figure 5-8 illustrates the P2: Abstract Symbolic Prediction by extending the line example from Section 5.1.2. During perceptual processing, a more amodal simulation that falls within a given abstract conceptual category (right line) can be engendered by a range of more concrete structures that could be emulated from the physical environment (left line). For example, many different emulated concrete structures (such as different trees in a physical environment) could engender simulations that fall within the abstract conceptual category ‘tree.’ Similarly, via a given graphic, when symbolic information conveys an abstract conceptual category with more certainty (right line), the symbolically conveyed abstract conceptual category applies to a range of concrete structures. In other words, when symbolic information increases (the certainty of symbolic information increases), uncertainty of the concrete structure conveyed decreases. For example, the written word ‘tree’ (right) symbolically conveys a more certain conceptual category within which many different more concrete examples (left) could fall.

Figure 5-8: In this demonstration for the Abstract Symbolic Prediction (Prediction 2), the written word ‘tree’ (right) symbolically conveys a more certain conceptual category within which many different pictorially represented more concrete examples (left) could fall.

I will formally demonstrate this prediction using the Euclid example from Part II and the programming language problem introduced in Chapter 1, but first let us first return briefly to the Chinatown example because it helps demonstrate the similarities and differences between pictorial information, on the one hand, and symbolic information, on the other: The English and Chinese sentences convey aspects, interpretations, or ‘abstract’ concepts about the food, such as ‘spicy,’ ‘hot,’ ‘sour,’ and ‘steaming.’ The same could be said for the visual programming
languages introduced in Chapter 1, which communicate concepts such as ‘infinity,’ ‘numbers,’ ‘parallel,’ and ‘square.’ A single picture can communicate concrete structures to people from various linguistic backgrounds because perceptual categories emerge similarly across cultures: the principles of occlusion, perspective, etc., can be found in English-speaking and Chinese-speaking cultures. The example of perceptually processing (pictorial information of) written graphics from a language that is unfamiliar to an audience helps demonstrate a role that pictorial information performs in engendering the construction of the simulations required for the symbolic information of written (more symbolic) graphics: Although emulated pictorial properties of the written Chinese graphics also use lower-level conceptual categories, enabling me to trace the outline of Chinese characters with my hand, I have developed capabilities to simulate the author’s intended symbolically represented ‘meaning.’

5.1.6 P3: Interference Prediction

Let us now consider the model’s third prediction, which anticipates that pictorial information interferes with symbolic information (when the author’s intention is to convey a more abstract concept). The third prediction builds upon the previous two predictions by focusing more explicitly on how they are interconnected, to anticipate that the cost of more pictorial information (and more concrete certainty) is less certainty regarding a conveyed abstract concept, which is manifested as less symbolic information.

How more pictorial information is predicted to correspond to less symbolic information. Recall Prediction 1, which anticipates that when more pictorial information is available, an (intended) concrete structure is conveyed with more certainty (Figure 5-5, left line), and that this (intended) concrete structure can be processed and filtered to fall within a range of possible abstract conceptual categories (Figure 5-5, right line). In other words, when more pictorial information is available, the certainty of the (intended) abstract conceptual category is reduced. This reduced conceptual certainty is the same as, or is akin to, less symbolic information. To illustrate this relationship, let us return to the apple example from Chapter 3. In Figure 5-9, a more pictorial graphic (a realistic picture of a piece of apple fruit) appears on the left. Compared with the text (“Apple Inc.”) on the right, the picture of the fruit is relatively more certain regarding the author’s intended concrete structure. For example, if the realistic picture
depicted a concrete structure from a physical environment (such as an apple), an artisan could probably use the realistic picture as a blueprint, or guide, to sculpt a concrete structure out of plaster and paint it to resemble the apple. However, the intended conceptual category (the ‘symbolic meaning’) intended by an author is less certain, because the depicted concrete structure could fall within a range of conceptual categories.

**How more symbolic information is predicted to correspond to less pictorial information.** Recall Prediction 2, which anticipates that when more symbolic information is available (conveys an abstract conceptual category with more certainty; the right line in Figure 5-8), the symbolically conveyed abstract conceptual category applies to a range of concrete structures. In other words, when symbolic information increases, the certainty of the concrete structure conveyed decreases. Because the conveyance of a concrete structure is the predicted affordance of pictorial information (Prediction 1), this decreased certainty about the graphic’s conveyed concrete structure would also seem to be a decrease in pictorial information. The written word “Apple Inc.” symbolically conveys a more certain conceptual category (more symbolic information) within which many different, more concrete examples could fall. The predicted relationship (or interference) between more pictorial information and less symbolic information can also be described in terms of limited resources, as discussed in the next section.

**Finite Resource Explanation: Drawing from Embodied Cognition**

As discussed in Chapters 2, 3, and the current Chapter 5, perception-reaction systems consist of more modally specific simulators, on one hand, and more amodal simulators, on the other (Kosslyn and Pomeranz 1977; Barsalou, 2009). Using the vernacular from Chapter 3 and the current Chapter 5, these more modally specific simulators are more closely coupled to the near ‘isomorphic’ emulation of light structures produced by a concrete physical environment. In contrast, more amodal simulation capabilities at higher-level convergence zones engender simulations that fall within more abstract conceptual categories. The key to the limited resource approach to describing interference is that simulation involves the same neural systems and capabilities required for perceptual processing of the physical world (Kosslyn and Pomeranz 1977; Barsalou, 2009). **Prediction:** Given that humans have a finite perceptual-cognitive capacity, when finite resources are consumed by processing pictorially represented, more modally specific, more concrete structures through emulation, those finite resources are less
available for the construction of more amodal simulations that fall within more abstract conceptual categories of symbolic information. In this way, pictorial information interferes with symbolic information by leaving fewer resources available for the construction of symbolic information. Pictorial properties are more problematic when the author’s intention is to engender the construction of audience simulations that fall within the author’s intended conceptual category.

Although it does not yet appear to have been applied to the graphic representation challenges addressed by this dissertation, recent work in the fields of embodied cognition and cognitive linguistics supports the principles that enable the foregoing prediction. Bergen referred to this as the ‘cross talk hypothesis’:

… the undisputed and recurrent finding, from a variety of methods, is that language about perceivable entities and events affects simultaneous or subsequent real perception, and that language about actions affects simultaneous or subsequent motor action…When people are conversing about topics that engage perceptual systems (language about visible or audible entities and events), or motor systems (language about performable motor actions), they might have a harder time perceiving the real world around them or performing the requisite actions involved in driving. (2012, p.5)

Bergen used empirical evidence to support his cross talk hypothesis, including a series of experiments that clarified why cell phone conversations, for example, cause so many distractions in the visual field while driving. Simply put, simulations engendered by (and required to understand) language use the visual-spatial cortex, which is also needed to emulate environmental features during driving. According to Bergen, “language with different content interferes more or less with higher-level perceptual reasoning and motor planning” (2012, p.17). Bergen’s use of the word ‘content’ is similar to the description of “an audience simulation that is intended by an author” in Chapter 3. Using the vernacular developed here, Bergen’s ‘content’ is an ‘intended percept’ (an intended concrete structure or abstract concept).

To illustrate Bergen’s ‘cross talk’ using an everyday example, consider the common experience of driving, searching for a desired exit, and simultaneously talking to passengers. It is not unusual for a driver to miss the exit, even if she was looking for the sign (see Bergen, 2012). Bergen’s suite of experiments suggest that this kind of increased interference is the result of simulation [of concrete structures] consuming the same resources and neural systems that would be required to emulate actual concrete structures through vision. Bergen wrote, “Visual
[concrete] language interferes with tactical control of a vehicle more than abstract language… because…its content [intended concrete structures or abstract concepts] can engage perceptual and motor systems vitally also deployed for perceiving the environment while driving and responding appropriately [words added for clarify relationship to the dissertation’s model]” (Bergen, 2012, p.17). These findings suggest that simulation required to process or generate spoken language is expected to interfere with emulation of an environment (and vice versa), but simulation of concrete structures engendered by spoken language interferes more with the emulation of concrete structures of an environment. These embodied cognition implications (ECI) could mean is that:

- **ECI 1.** When finite resources are occupied for the purposes of simulating concrete structures, fewer resources are available for:
  - **ECI 1b.** the emulation of concrete environmental structures other than those being simulated, and
  - **ECI 1c.** the simulation of abstract concepts.

- **ECI 2.** When finite perceptual-cognitive resources are occupied for the purposes of emulating concrete physical structures, fewer resources are available for:
  - **ECI 2b.** the simulation of concrete structures other than what is being emulated, and
  - **ECI 2c.** the simulation of abstract concepts.

I make a similar argument via Prediction 3: that these phenomena also appear to help explain affordances of external graphic representations (i.e., when emulation of a more modal concrete structure interferes with the processing of symbolic properties and vice versa). You have probably experienced this phenomenon in very tangible ways in everyday life, trying to process spoken (symbolic) language via auditory information while simultaneously processing (symbolic) information via visual information. Perhaps you have been at a café trying to read or write while a person at a nearby table speaks loudly on a cell phone. For many people, perceptually processing the conversation impedes their ability to read or write text; a leading hypothesis that predates the embodied cognition view is that interference is caused because auditory processing of spoken language consumes resources required to simulate the ‘inner voice’ needed to read or write (c.f., Salami & Baddeley, 2004). However, the embodied cognition noted previously goes a step further, by suggesting that interference is also caused
because language understanding requires simulated ‘reenactments’ that recruit perceptual processing capabilities required for perceiving and reacting within the physical world.

Let us now consider Prediction 3 by further extending the line example. Figure 5-9 integrates the components of Figures 5-5 and 5-8. Figure 5-5 focused on the predicted affordances of a more pictorial graphic (where more information is available about a concrete structure and less information is available about an abstraction) and Figure 5-8 focused on the predicted affordances of a more symbolic graphic (where more information is available about a concrete structure and less information is available about an abstraction). Figure 5-9 focuses on a spectrum of graphics, from more pictorial (left) to more symbolic (right); each vertical line corresponds to a hypothetical more pictorial or more symbolic graphic. It helps show how more pictorial information (of a given graphic) corresponds to less symbolic information, and vice versa. Figure 5-5 would correspond to a vertical line on the left of Figure 5-9; Figure 5-8 would correspond to a vertical line on the right of Figure 5-9.

In Figure 5-9, the height of each vertical line indicates how perceptual-cognitive capacity is constrained, finite, and limited. Each point on the line corresponds to a possible concrete structure, on one hand, or possible abstraction, on the other. This constraint-limitation means that when, for example, a graphic is highly pictorial, emulation resources are consumed through the processing of a depicted concrete structure, shown by the highlighted blue segment of a line. This emulated concrete structure can fall within a range of more abstract conceptual categories, as shown by the red highlighted segment of a line. Because resources are consumed by the emulation of a concrete structure, fewer resources are available for more amodal simulations that could fall within a more certain abstract conceptual category.

In a ‘more pictorial’ graphic, where pictorial information conveys a concrete structure with more certainty, emulation of the graphic that was configured by an author engenders pictorial information during perceptual processing. The shorter blue highlighted segments of the lines to the left in Figure 5-9 indicate this pictorial certainty (more information about an intended concrete structure). The red highlighted segment of the line indicates how the emulated concrete structure could be processed and filtered to fall within a range of possible more abstract conceptual categories, and is therefore predicted to convey less symbolic information. In this
way, an increase in pictorial information is predicted to correspond with a decrease in symbolic information.

Conversely, in a ‘more symbolic’ graphic, where symbolic information conveys an abstract concept with more certainty (the shorter red highlighted lines to the right of Figure 5-9), a concrete structure is conveyed with less certainty, because numerous concrete structures could fall within the symbolically conveyed conceptual category. This reduced certainty about the conveyed concrete structures is indicated by the longer blue highlighted lines to the right of Figure 5-9). In this way, an increase in symbolic information is predicted to correspond to a decrease in pictorial information.

Figure 5-9: The Interference Prediction (Prediction 3) anticipates that the cost of more pictorial information (and more concrete certainty) is less certainty regarding a conveyed abstract concept, which is manifested as less symbolic information.

Certain aspects of Prediction 3, which will be detailed in the second half of this dissertation, can help clarify the prediction in more detail. For example, suppose an author intends to convey an abstract concept (Intended Abstraction B in Figure 5-10), but miscommunicates by selecting a more pictorial graphic (Pictorial Graphic A in Figure 5-10). Pictorial Graphic A is tasked with communicating an Intended Abstraction B (such as the abstract concept of a rectangle in a geometric proof) but could unintentionally draw attention to
an unintended concrete instantiation of the Intended Abstraction B (such as the specific concrete square depicted by Pictorial Graphic A). The price of this inefficiency can be described as an increased use of our finite simulation capacities.

Figure 5-10: An example of miscommunication caused by a graphic.

Reducing interference. For an author to engender the construction of an intended audience simulation with more certainty (i.e., endeavour to be precise), she must use culturally constructed capabilities to use symbolic properties. As described in Chapter 3, Claim 2, capabilities to construct Symbolic Properties are developed through the (often systematic) construction of shared intentionality and conventions among individuals, often through standards that are maintained (and/or enforced) by cultural institutions. These capabilities enable an author
to simulate how an audience might perceptually process a graphic, employing the recursive process of shared intentionality described in Chapter 3.\textsuperscript{37}

**Role of shared intentionality in reducing symbolic uncertainty.** To use everyday vernacular, shared intentionality enables both an author and their audience to be more certain about what ‘symbolic meaning’ was conveyed and received. In other words, for a given configuration of marks, shared intentionality should reduce the range of possible audience interpretations that an author might need to anticipate (Figure 5-11, left). Additionally, shared intentionality allows different configurations, or types, of marks to correspond to expected symbolic meanings.

\textsuperscript{37} ‘Shared expectations’ or ‘shared understanding’ are other ways of referring to the process of shared intentionality described in Chapter 3.
Figure 5-11: Shared intentionality enables an author to anticipate how an audience might process pictorial and symbolic properties to receive intended concrete structures or abstract concepts. It reduces the range of possible audiences emulations that an author is required to simulate.

5.1.7 Pictorial and Symbolic Information: Some Conclusions

This chapter applied concepts from the preceding sections and chapters, using theories about perceptual and conceptual categories to predict affordances of pictorial and symbolic information:

**P1: Concrete Pictorial Prediction.** Pictorial information affords the communication of concrete concepts (more effectively than symbolic information).

**P2: Abstract Symbolic Prediction.** Symbolic information affords the communication of abstract concepts (more effectively than pictorial information).

**P3: Interference Prediction.** Pictorial information interferes with the communication of the abstract concepts of symbolic information.
This chapter developed these predictions by introducing the well-researched concept of ‘attention’ to this dissertation's conceptualization of emulation versus simulation, thereby predicting affordances of pictorial and symbolic information. The significance of this chapter is its introduction of a predictive theory that can be tested and demonstrated, which will be accomplished via a site of artefact evolution in the next chapters.

5.2 Predicted Affordances of Pictures, Diagrams, and Sentences

To predict affordances of pictures, diagrams, and sentences, I can apply the P1: Concrete Pictorial Prediction, the P2: Abstract Symbolic Prediction, and the P3: Interference Prediction to the previous conceptualization of pictorially and symbolically represented objects and relations (C4: Diagrammatic, C5: Sentential, and C3: Composite Pictorial Claims) to understand how each type of graphics affords different kinds of actions. Simply, if pictorial properties afford the communication of concrete concepts (because they are more perceptually/modally specific) whereas symbolic properties afford the communication of abstract concepts (because they are less perceptually/modally specific), then the following predictions result:

**P5: Diagrammatic Prediction.** Diagrams, as pictorially represented relations among symbolically represented objects (C4: Diagrammatic Claim), are predicted to afford the communication of concrete relations among abstract concepts.

**P6: Sentential Prediction.** Sentences, as symbolically represented relations among symbolically represented objects (C5: Sentential Claim), are predicted to afford the communication of abstract relations among abstract concepts.

**P4: Composite Picture Prediction.** Pictures, as pictorially represented relations among pictorially represented objects (C3: Composite Picture Claim), are predicted to afford the communication of concrete relations among concrete concepts.

The next task is to demonstrate these predictions by clarifying how logical structures (such as proofs or computer programming languages) have been naturally selected as text-sentential paradigms over time because the text-sentences of proofs or textual codes of programming languages afford the communication of abstract relations among abstract objects.
Additionally, pictures (because of their concrete relations among concrete objects) and diagrams (because of their concrete relations among abstract objects) afford the design of logical structures such as proofs or computer programs.

5.3 Demonstrating Predictions

One strategy for supporting, demonstrating, or 'testing' claims/contributions is to show how predicted affordances correspond to functional purposes. I will begin by demonstrating the P2: Abstract Symbolic Prediction. According to this prediction, symbolic properties should be more effective than pictorial properties at communicating abstract concepts (by communicating abstract concepts with more conceptual certainty), when the functional purpose of the graphic is to communicate an author's intended abstract concept. However, the following methodological concerns must be addressed in order to demonstrate the P2: Abstract Symbolic Prediction:

a. First, it is necessary to find a convincing way to demonstrate how a clearly identifiable type of graphic corresponds to a functional purpose.

b. Second, it is necessary to find a way to demonstrate that a graphic is effective relative to a functional purpose.

Here, I turn to the principals of artefact evolution discussed previously. In artefact evolution, we can assume that principles of natural selection will cull any practices and conventions that are less effective relative to a functional purpose, leading to the emergence of practices and conventions that are more effective relative to a functional purpose. For artefact evolution to be a convincing method, however, I must also demonstrate that:

c. Selection pressures are sufficiently high to encourage the natural selection of artefacts to somewhat effective states or configurations relative to functional purposes.

Artefact evolution can be used to demonstrate the explanative prediction that when the purpose is to communicate an abstract concept, symbolic properties will be more effective than pictorial properties (by communicating abstract concepts with greater certainty). Based on the principles of artefact evolution, *symbolic properties are more likely than pictorial properties to be naturally selected, when the functional purpose is to communicate abstract concepts*. I will refer to this prediction as the Behavioural Symbolic Information Prediction (BSIP).
Now I can turn to the natural selection of the sentence-based axiomatic method, using Euclid's *Elements* as a site of artefact evolution. Within this site of artefact evolution, artefacts from a long period of artefact evolution are observable. I will trace how Euclid's *Elements*, the first known example of how what is commonly referred to as the ‘axiomatic method’ (which is generally considered to be sententially-based), was naturally selected from a 'primordial soup' of pictures. I will use this example to demonstrate that sentences were naturally selected over pictures when the functional purpose was to afford the communication of abstract concepts, which suggests that sentences are a more effective way to communicate abstract concepts. In other words, I will use the observable evolutionary history from pictures to sentences, and will map the functional purposes of particular pictorial or symbolic graphic types during each era; this will illustrate BSIP, as an indirect way of demonstrating the model's P2: Abstract Symbolic Prediction.

Along with these evolutionary observations, I will also refer to philosophical debates about the controversial role of pictures during key phases of that evolution. The so-called ‘arithimatization’ of mathematics in the 19th Century was controversial; scholars asked why pictures were considered to interfere more than sentences with the communication of abstract concepts. Because sentences were naturally selected over pictures (and in earlier eras, these pictures afforded ancient land surveying tasks), I will be able to illustrate the prediction that *pictures afford the communication of concrete concepts* (P1: Concrete Pictorial Prediction). I will refer to this as Behavioural Pictorial Information Prediction (BPIP). Finally, because a number of mathematicians, logicians, and interface designers have claimed that diagrams 'reduce inferential load' during logical reasoning, I will identify similarities between the roles played by diagrams during intermediate phases of software design and the evolution of the axiomatic method to explain, via the model, how diagrams afford intermediate phases of design.
Figure 5-12: Pictorial and symbolic information mapped to abstract and concrete concepts.
Part II: Demonstrating the Model
6 Demo #1: Why are (Geometric) Proofs ‘Non-Visual’?

6.1 Introduction

To this point, this dissertation has described a gap in the current theoretical understanding of the properties (and affordances) of graphic representations typically used in the design of visual information displays. It also proposed a novel method for engaging in observations to respond to this gap, and then outlined a theoretical model to fill the gap. Next, it will demonstrate the model's usefulness when applied to a specific graphic representation phenomenon.

The next sections will apply the model to describe the affordances of graphic representations that have evolved through the process of artefact evolution. This description will be used as a step toward predicting graphic representations more widely.

This subsection is an important transition in this dissertation. Previous chapters and sections developed a general model for describing the properties of graphic representations; here, I begin to describe graphic representations in a specific context, and use this description and context to formulate predictions about graphic representation affordances more widely. The evolutionary history of the axiomatic method can help clarify the properties of contemporary human–computer interfaces, and programming languages in particular.

**Problem.** The problem being addressed here is clarifying why geometric proofs are ‘non-visual,’ and specifically why they are composed of sentences. The answers will help elucidate a larger problem: why programming languages are ‘non-visual’ and sententially based. The results should help explain why most ‘visual’ programming languages have failed, and what it means for an interface to be ‘intuitive.’ Most contemporary text-based programming languages are based on the axiomatic method (cf. Hoare, 1969; 1999). The axiomatic method is considered the foundation for the scientific method (Betti & De Jong, 2010) that is used in many of the natural sciences, engineering, and social sciences.

...
Why are most geometric proofs usually ‘non-visual,’ i.e., textually based? The subject matter of geometry and its underlying theories intuitively involves spatial forms and relationships, but geometric proofs are usually formally represented as text-based descriptions of geometric properties, which demonstrate how a geometric relationship is necessarily true via a series of logical relationships. Tennant (1986, p. 304) wrote: “[The diagram] is only an heuristic to prompt certain trains of inference; . . . it is dispensable as a proof-theoretic device; indeed, . . . it has no proper place in a proof as such. For the proof is a syntactic object consisting only of sentences arranged in a finite and inspectable array” (as quoted in Barwise & Etchemendy, 1991, p.3).

For example, if block A is below block B and block C is above block B, then logic tells us that block A is below block C. We can easily infer that block A is below block C by looking at a diagram, but the logical text-based proof is considered more rigorous than a diagram (Tennant, 1986). For generations, Euclid’s *Elements* was considered to be flawed because of its reliance on diagrams. According to Mumma, “for some of Euclid’s steps, the logical form of the preceding sentences is not enough to ground the steps. One must consult the diagram to understand what justifies it” (2009, p.222). For this reason, most modern scholars feel that Euclid “failed in his efforts to produce (an) exact, full explicit mathematical proof” (Mumma, 2008).

In 1991, Barwise and Etchemedy questioned the assumption that diagrams were less rigorous than (non-diagrammatic) proofs, and suggested that diagrams might actually offer advantages. They strengthened their argument using evidence drawn from the field of cognitive psychology (Kosslyn, 1980), which demonstrated that in certain situations, maps enable more efficient problem-solving. Nonetheless, most mathematicians still prefer text-based ‘language thinking’ and algebraic notations to diagrams (Brown, 1999). So, the question here is: Can text afford ergonomic properties better than diagrams, and thereby encourage textual modes for proofs?

Before going much, it is important to review some distinctions between formal and informal proofs and between different cognitive tasks such as proof generation, presentation, and comprehension. Because this discussion focuses on features that distinguish the comprehensibility of diagrams from a human cognitive perspective, we are more interested in informal proofs, because formal proofs are not primarily designed for human use. Furthermore,
we are more interested in diagrams as communication tools to present and communicate proofs, rather than in their use as primary reasoning tools to generate proofs.

**Chapter thesis.** The central thesis presented in this chapter applies the model’s ability to predict affordances, as outlined in the previous chapter. Specifically, it focuses on demonstrating the P2: Abstract Symbol Prediction, according to which symbolic information affords the communication of abstract concepts with more certainty than pictorial information. Additionally, this chapter focuses on demonstrating the related P3: Interference Prediction, according to which pictorial information interferes with the construction of simulated prototypes (which involve an author's intended abstract conceptual categories), and therefore interfere with the processing of symbolic information.

Within the context of this site of artefact evolution, the P2: Abstract Symbol Prediction suggests that when the functional purpose is to communicate an abstract concept, symbolic properties should be more effective than pictorial properties, because they can communicate abstract concepts with more conceptual certainty. Based on the principles of artefact evolution, when the functional purpose is to communicate an abstract concept, symbolic properties are therefore more likely to be naturally selected over pictorial properties (Observable Behavioural Prediction).

The main goal of this chapter is to demonstrate the P2: Abstract Symbolic and P3: Interference Predictions, but this chapter will also help demonstrate the Concrete pictorial Prediction, because while sentences were naturally selected over pictures within this natural site of artefact evolution, pictures clearly serve a concrete functional purpose.

**Chapter Significance.** This chapter makes several significant contributions: 1) it demonstrates key aspects of the model’s claims and predictions; 2) it explains why geometric proofs may have been naturally selected to their current ‘non-visual’ sentential form; 3) it identifies clues about why most programming languages are sententially based; and 4) it establishes the first parts of a predictive model that can be used to develop a science of visual information design.

**Scope and limits.** Many kinds of graphic representations currently exist and could be explained using the proposed model, so here it is useful to clarify the scope and limitations of the
demonstrations that follow. This chapter will focus primarily on the natural selection of sentential and pictorial graphic representations in a specific example of one of Euclid’s propositions, Proposition 35 of Euclid’s *Elements* (Appendix I) – a proposition that has been discussed by others in the context of the questions here. In particular, it will explore why sentences were naturally selected over pictures, and why the pictures that were naturally selected were so controversial. Using the vernacular of Gregor’s (2006) taxonomy of theories, it will develop a Type II theory (explanation, not prediction). Furthermore, it will demonstrate the potential predictive power of the model by developing a Type III-IV theory about predicting the perceptual-cognitive affordances of sentences in communicating logical structures.

The uptake and widespread acceptance of a notation and inference system involves various social and psychological features, so our theory can help to explain why some systems appeal to mathematicians more than others in the context of specific tasks.

6.2 History of a Site of Artefact Evolution

This subsection discusses why geometric proofs, particularly Proposition 35 of Euclid’s *Elements*, can serve as a site of artefact evolution (SAE) to demonstrate how the properties of symbolic and pictorial information afford perceptual-cognitive reactions.

The axiomatic method has humble origins: geometry (meaning ‘land measurement’) evolved from land surveying and agricultural activities in many early cultures, including ancient Babylon (see Figure 6-1; Kneale & Kneale, 1962, p. 2). Early proofs were distinctly phenomenological, and even pictorial:
Figure 6-1: Babylonian clay tablet YBC 7289 with annotations. The diagonal displays an approximation of the square root of 2 in four sexagesimal figures, to approximately six decimal figures. Courtesy of Bill Casselman (http://www.math.ubc.ca/~cass/Euclid/ybc/ybc.html) and the Yale Babylonian Collection.

The earliest version of mathematics was phenomenological. The only justification required for a mathematical ‘fact’ was a plausible picture or a compelling description (sometimes using analogy, or even by invoking the gods). The idea that mathematical statements could be proven had not yet been developed: there was no standard for the concept of proof, and the logical structure (the “rules of the game”) had not yet been created (Krants, 2007, p. 2).

Scholars believe that the transformation from this humble phenomenological form to the axiomatic method occurred in ancient Greece around 300 BCE (Kneale & Kneale, 1962; Krants, 2007). Although commonly attributed to Euclid, the work presented in the *Elements* is actually a synthesis of previous research spanning multiple generations and cultures (Heath, 1908, p. 29). Earlier mathematicians had developed proofs, including Hippocrates of Chios, Eudoxus, and other Pythagorean and Athenian mathematicians (Ball, 1908, p. 54). Thales (624–546 BCE) is
thought to have proven some geometrical theorems. Others, such as Eudoxus (408–355 BCE) and Theaetetus (417–369 BCE), are credited with formulating theorems but not proving them. Aristotle (384–322 BCE) described how definitions should be developed from familiar concepts (Ball, 1908).

By 300 BCE, Euclid was able to synthesize this previous work and introduced the axiomatic method that is still being used: start with undefined terms and axioms (propositions that are assumed to be self-evident) and deductively apply those axioms using at least one sound rule of deduction/derivation to prove theorems (Heath, 1956). The *Elements* played a central role in Western education until the middle of the 20th century (Boyer, 1991), when the critical role afforded by pictures was identified as a flaw (Mumma, 2010).

In summary, while the work presented in the *Elements* is often attributed solely to Euclid, it is really the product of a long ‘bottom-up’ evolution of what Dawkins (1976) called ‘memetic structures’ – units of culture, e.g., an idea that is ‘hosted’ in many minds. The work Euclid synthesized has withstood the test of time, and this chapter will explore how Euclid’s synthesis has since been further ‘naturally selected’ into subtle variations, including some used in contemporary mathematics and programming languages. Although some may argue that the axiomatic paradigm is merely convention, this chapter will demonstrate that the definitions, postulates, and axioms that emerged from Euclid’s work also ergonomically afford certain kinds of perceptual-cognitive tasks.

Similar to how the architecture of human intelligence and creativity can be elucidated based on its evolutionary history, this chapter will explore the structure of modern interface obstacles based on evolutionary history. One additional factor makes this site of artefact evolution even more useful: the fact that Euclid’s use of pictures became controversial over time. The next section focuses on this controversy and how it points to differences between the properties of pictures and text.

### 6.3 Controversy Emerges in the 19th Century…Due to Pictures

The *Elements* was considered the exemplar of deductive argument and was used as the primary textbook for mathematics until the middle of the 20th century (Boyer, 1991). However, the pictorial aspects of the *Elements* became controversial in the 19th century. Frege and others
grew increasingly concerned that the role afforded by these pictures left mathematics on unstable ground (Mumma, 2009). In short, scholars now consider Euclid’s reliance on pictures to disqualify his proofs as rigorous because they cannot ‘stand on their own’ mathematically and must be supplemented by additional axioms. According to Mumma:

For most of its long history, Euclid’s *Elements* was the paradigm for careful and exact mathematical reasoning. In the past century, however, it has been just the opposite. Its proofs are often invoked to illustrate what rigor in mathematics does not consist in. Though some steps of Euclid’s proofs are respectable as logical inferences, a good many are not. With these, one cannot look only at the logical form of the claims in the proof and understand what underlies the inference. One is forced, rather, to look at the accompanying diagram. The modern opinion is that Euclid’s proofs exhibit a deductive gap at such places. (Mumma, 2009, p. 2)

In the late 19th century, Hilbert addressed this problem by assembling a text-sentential axiomatic model on which modern mathematics could be firmly based. This was considered a revolutionary work, and most mathematicians now consider his model to be more stable than Euclid’s.

Hilbert’s work placed Euclid’s theorems within a mathematical context where all modes of inference were laid out explicitly in advance. This development is now universally regarded as a methodological breakthrough: geometric relations that were once logically free-floating, because they were presented via diagrams, were given a firm footing with precisely defined primitives and axioms (Mumma, 2008).

However, dissenters emerged about 20 years ago, arguing that diagrams and other so-called heterogenous reasoning systems were not only valid, but sometimes superior. The next section discusses the emergence of a small subfield: heterogenous and diagrammatic reasoning.

### 6.4 A Defense of Pictures Emerges: Heterogenous and Diagrammatic Reasoning

From an information-processing (cognitive psychology) perspective, Larkin and Simon (1987) explored the differences between information as diagrams versus information as sentences, concluding that sentences embody the characteristic of being indexed on a list, with each element ‘adjacent’ only to the next element in the list. In contrast, diagrams are indexed by location on a plane: many elements may share the same location, and each element may be
adjacent to any number of other elements. For this reason, Larkin and Simon proposed that diagrams are more useful than sentences for solving certain kinds of problems because they can support more efficient computational (including human neurological) processes. They also noted that this efficiency depends on the design of the diagram and the ability of the user to interpret the diagram.

Approaching the same issue from a different perspective, Barwise and Etchemendy began by noting that within the fields of mathematics and logic, diagrams are not considered valid parts of a proof; they only appear as heuristic aids (Barwise & Etchemendy, 1991, citing Tennant, 1986). Their goal was to change this perception, bolstering their case by citing work by Kosslyn (1980), a cognitive psychologist who used maps to justify visual presentations as valid problem-solving tools and argued that sentences or visual representations offer advantages or disadvantages based on the purpose of the task at hand.

Barwise and Etchemendy also concluded that diagrams provide advantages that are not offered by sentences. They extended on Larkin and Simon’s work, identifying various differences between sentences and diagrams. For example, relationships are often implicit in diagrams, whereas even the most trivial consequences must be inferred explicitly using sentences (as demonstrated in the Introduction). They also argued that a picture or diagram can support ‘countless facts’ (i.e., a plurality of sentences).

Shimojima and Katigiri (2008) supported the work done by Barwise and Etchemendy (1991) by collecting empirical evidence to prove that diagrams reduce inferential load, drawing on modern cognitive science research such as Ballard et al.’s (1990) theory of ‘deictic indices.’ Deictic indices act as mental ‘pointers’ to particular objects in external space. According to the theory, this attentional mechanism enables individuals to maintain a small pool of such indices simultaneously and easily direct their focal (mental) attention or gaze to any of these indices. In other words, “pointing movements are used to bind objects in the world to cognitive programs” (Ballard et al., 1990, p. 723)

Shimojima and Katigiri (2008) described how these indices can be used to keep track of (mental) ‘non-physical drawings’ constructed by individuals as they make inferences about diagrams. They suggested that individuals navigate through these non-physical (mental)
drawings during reasoning tasks; the fact that these non-physical drawings are assisted by ‘real’
drawings (diagrams) reduces the ‘inferential load’ required during reasoning tasks.

Although none of these works specifically address our question about why proofs are
‘non-visual,’ they provide important insights about the differences between text and diagrams,
which can help guide our inquiry. The common thread running through all of the research
described above is that text/language/prose guides attention in ways that are different from
visual-spatial representations.

6.5 Extrapolation: Does Language Guide Attention through
Visual-Spatial Structures?

From this loose collection of interrelated work, we can extrapolate some general
principles regarding the cognitive affordances of ‘visuals’ relative to text (Coppin & Hockema,
2009; Coppin, Burton, & Hockema, 2010). As discussed above, Larkin and Simon (1987)
suggested that sentences have a list-like structure, in that each item on the list is only adjacent to
the item before or after it on the list. In contrast, items in a diagram are adjacent to many items
on a list. This approach is similar to that taken by Barwise and Etchemendy, who suggested that
a picture or diagram can support ‘countless facts’ (i.e., a plurality of sentences can be constructed
from a diagram). In other words, many sentences could be created by linking together elements
of a diagram into sentences (list-like structure). In this way, we can extrapolate that a list-like
structure (as suggested by Larkin and Simon) can be inferred from a diagram. Each sentence
inferred from a diagram is like a path that guides attention through the visual-spatial
relationships in a diagram. Shimojima and Katigiri demonstrated this extrapolation using an eye-
tracking study, which revealed reasoners mentally guide their attention through a “non-physical
drawing” (2008, p. 87). They suggested that actual drawings support non-physical (mental)
drawings, thereby reducing inferential load. In summary, sentences appear to guide attentional
paths through both physical and non-physical (mental) visual-spatial structures. Further,
Shimojima and Katigiri appear to have demonstrated that rational language/propositional logic
guides attention and motor movements (through eye fixations) through non-physical visual-
spatial representations.
6.6 An Attention-Based Theory to Explore Affordances of Textual and Diagrammatic Proofs

Expanding on prior research about attention, my colleagues and I proposed that textual graphics may focus a reasoner’s ‘spotlight’ of attention through serialized sequential chunks, enabling the methodical presentation of a rational argument in a way that is not possible (or at least very difficult given current understandings and practices) when using a diagrammatic notation that may ‘diffuse’ attention. Diagrammatic notations may enable a reasoner to discern how elements fit together holistically, but provide less clarity about individual steps in the argument, or about reifying and explicitly representing the connections/relationships between steps (Coppin & Hockema, 2009; Coppin, Burton, & Hockema, 2010).

In these previous studies, we drew upon modern theories of attention, such as Treisman (1980), who focused on focused attention (FA), suggesting that “attention must be directed serially to each stimulus in a display whenever conjunctions of more than one separable feature are needed to characterize or distinguish the possible objects presented” (1980, p. 97). By describing features as separable, she was referring to primitives (such as basic shapes and colors) prior to integration into a conceptual whole. She used a spotlight/zoom lens metaphor, arguing that attention can be narrowed to focus on a single feature, when an individual needs to see what other features are present and mentally form an object, or diffused over a whole group of items that share a relevant feature. Compared with diffused attention (DA), FA is more precise and detailed. In our previous research, my colleagues and I argued that textual language guides FA through such structures to build more precisely focused but less holistic ideas, thereby helping to explain the origins of an experience of rigor (Coppin, Burton, and Hockema, 2010).

Some scholars have argued that so-called ‘object-based’ attention may automatically (‘pre-attentively’) spread as a ‘bottom-up’ process within ‘groups’ of objects, but that specific queries and other factors can influence a ‘top-down’ process that changes which groups are perceived as objects (Scholl, 2001). For example, Scholl described how an initial Gestalt grouping of two intersecting lines can change to a different Gestalt grouping based on a statement (e.g., ‘identify the bird beaks in a field of intersecting lines’). A text description in a proof (e.g., ‘consider vertices A-B-C-D’) could cause an individual to perceive a quadrilateral within a field of many intersecting lines. This process is relevant to our usage of FA and DA, and helps develop the following explanation for why proofs are generally ‘non-visual.’
Why a diagram cannot usually constitute a convincing ‘holistic’ proof: The need for symbols to fall within the narrow spotlight of focal attention means that diagrams, and the spatial relationships they embody, cannot usually be attended to all at once. Therefore, they must be processed in a way that allows linkages between earlier perceptual memories and later percepts.

Why text is effective for proofs: External symbolic representations, such as text, enable each symbol to fall sequentially within the narrow focus of focal attention.

Why proofs are often sequential: Focal attention must move through various parts of a conceptual visual-spatial structure due to narrower, but more intense, focus (rather than attending to the whole structure simultaneously).  

Why propositional logic is used for proofs: Propositional structures in proofs may provide symbolic shortcuts, serving as stand-ins for visual-spatial relationships that cannot all simultaneously be in the narrow spotlight of focal attention. For example, the statement “if C is below B” references a perceptual symbol constructed from a previously considered image, as does the statement “if B is below A.” The conclusion “therefore C is below A” references the two previous symbols, and supports the construction of a new mental image that can serve as the basis for a new perceptual symbol and be used in future propositional statements.

6.7 Mumma’s Defense of the Elements via Exact and Co-exact Properties

Mumma (2009) recently explored how the methodology of Euclid’s Elements differs from that of modern theories of geometry, characterizing Euclid’s proofs as essentially diagrammatic. Mumma argued that Euclid developed a formal system with a ‘diagrammatic symbolic type,’ countering the standard modern assessment of Euclid’s methodology as intuitive and un-rigorous. In so doing, Mumma identified three interrelated factors that made scholars consider Euclid’s reliance on diagrams in his proofs to be non-rigorous:

38 It should be emphasized that we are using the notion of saccade here metaphorically; sequential movement of attention to various parts of a structure will probably bear no resemblance to the way an eye actually saccades, such as to its ballistic nature for example. Further, the “sequence” implied by the word “sequential” here is not necessarily imply a particular ordering, especially not one that might correspond with the actual sequence of eye saccades. We assume that many cognitive and pragmatic factors play into determining in what order structures must be attended.
• The generality problem – proofs are meant to be more general than the particular instance in a diagram, but it is difficult to generalize from a particular diagram to a more general case.

• The modern mathematical understanding of continuity – diagrams may lead to simplistic and invalid assumptions about the continuity of lines, e.g., with respect to the existence of intersection points.

• The modern axiomatic method – all axioms and deductive steps must be explicitly specified, which raises suspicion about the assumptions embedded in diagrams.

Mumma (2009) then argued that the three problematic factors can be overcome in a more carefully specified hybrid system. He began by introducing two terms: exact properties and co-exact properties, and explained this distinction as follows:

Any one of Euclid’s diagrams contains a collection of spatially related magnitudes—e.g. lengths, angles, areas. For any two magnitudes of the same type, one will be greater than another, or they will be equal. These relations comprise the exact properties of the diagram. (Mumma, 2009, p. 10)

He also wrote:

How these magnitudes relate topologically to one another—i.e. the regions they define, the containment relations between these regions—comprise the diagram’s co-exact properties. (Mumma, 2009, p. 10)

In a subsequent work, Mumma provided an example:

Diagrams of a single triangle, for instance, vary with respect to their exact properties. That is, the lengths of the sides, the size of the angles, the area enclosed, vary. Yet with respect to their co-exact properties the diagrams are all the same. Each consists of three bounded linear regions, which together define an area (Mumma, 2010, p. 223).

Mumma’s (2009) analysis thus seems very helpful to answer our motivating question here: if diagrams are more efficient (in terms of inferential load) than text in reasoning tasks, why were text-sentential graphic representations naturally selected for communicating proofs? Mumma would say the reason is because mathematicians do not find them rigorous enough for
the reasons listed above, all of which are rooted in the subject matter and norms of the field. While this is persuasive, it should be noted that in each of Mumma’s three reasons given above, there are embedded and unexamined assumptions about the limitations of diagrams for communicating. For example, consider the following phrases from the bullet points on the previous page:

- … but it is difficult to generalize from a particular diagram …
- … diagrams may lead to simplistic and invalid assumptions …
- … which raises suspicion about the assumptions embedded in diagrams [that they are somehow not explicit].

Each of these represent assumptions about diagrams that may or may not be valid. All three of these beg the question “why”. For example, regarding the first, in order to make the case that no diagrams could effectively support generalization, we must understand why it is difficult to generalize from them. Thus, while Mumma’s explanation for mathematicians’ general sense that diagrams lack rigor may be appealing at one level, at a deeper level it calls out a need for the model that this dissertation is trying to develop. Finding an answer to these questions could help clarify the differences between the properties of text-sentences and the properties of pictures. The next sections will try to apply some of Mumma’s (2009) helpful distinction between exact and co-exact properties to demonstrate how exact properties are akin to pictorial information in my model, whereas co-exact properties are akin to symbolic information.

6.8 This Model's Response to the Controversy

6.8.1 Some Notes about ‘Rigor’

Before presenting our theory, it is important to define what is meant by a user’s relative assessment of ‘rigor’ regarding two mathematical notations. The rigor of a certain notation refers to a mathematical quality, which is either present or missing. In the context of this work, the notion of rigor is used to quantify the clarity and persuasiveness of a notation, resulting in a user’s greater or lesser degree of confidence in the notation. In this context, users are mathematicians, who use various notations to communicate with colleagues via proofs. The assessment of rigor may be related to, but separate from, logical soundness. A mathematician may use a visual notation with a formal semantics as an integral part of a proof, with confidence
that the results are sound. If a formal diagrammatic element of a proof fails to inspire a feeling of rigor in readers (perhaps because the notation uses semantics that are complex and unfamiliar), the proof may have failed to fulfil its purpose as a communication tool. Conversely, the same mathematician may use an informal notation as an aid to reasoning with confidence, when she expects the meaning will be clear in the context it is used. The theory being developed in this dissertation aims to provide a partial explanation of the clarity and confidence experienced by users of textual proofs, and to extract features that can be designed, primarily in visual notations.

6.8.2 The Model’s Response to the Controversy

Why are geometric proofs non-pictorial? The P2: Abstract Symbol Prediction suggests that if an author wants to communicate an abstract concept, then pictured objects, relations, or attributes would not effectively afford the communication of those intended abstract objects, relations, or attributes (due to the conceptual ambiguity of pictorial properties). Additionally, based on the P3: Interference Prediction, pictorial properties would impede the communication of intended abstract concepts. According to the model developed here, the reason for this is that concrete concepts/structures use emulation capabilities (lower-level perceptual categories), which are closely linked with sensory receptors, and developed via perception-reactions with the physical world.

Instead, symbolic properties would be more effective at communicating abstract concepts, following the P2: Abstract Symbol Prediction. Due to the principles of artefact evolution, symbolic properties, and symbolized objects, attributes, and relations would be naturally selected over time, forcing pictorial conventions out of favour (BSI Prediction). This is what we see in this site of artefact evolution.

6.8.2.1 Explaining the Natural Selection of Sentences from a ‘Primordial Soup’ of Pictures

Ancient land-surveying techniques, from which the axiomatic method and Euclid’s *Elements* were naturally selected, were apparently afforded by outline drawings (Section 6.2; Kneale & Kneale, 1962; Krants, 2007; Figure 6-1; Figure 6-2, Label A). This allows the emulation of a given physical shape of surveyed land, using lower-level perceptual categories that are more closely linked with sensory receptors of vision, and that emerged while perceiving and acting within a physical environment composed of occluded surfaces and edges (Chapter 3,
Chapter 5). So, employing the vernacular introduced thus far, when a (non-graphically represented) physical land structure is perceptually processed, it is concrete (Chapter 5). If an author wants to communicate the physical structure of that land, she can do this with a degree of perceptual certainty using a pictorial graphic (P1: Concrete Pictorial Prediction) that targets emulation capabilities and associated lower-level perceptual categories to generate the intended emulation: pictured occluded surfaces and edges that are beyond, or other than, the marked surface of the picture.

Figure 6-2: Natural selection of sentences from a primordial soup of pictures.

Over time, more abstract mathematical principles were discovered or developed by philosophers, mathematicians, and logicians (Section 6.3; Mumma, 2008; Figure 6-2, Part B). According to Barsalou's account of abstraction, these were less modally specific/more amodally simulated prototypes that can be classified as more abstract conceptual categories, which are generated via the integration or convergence of perceptual modes at higher-level association areas (see Chapter 5). As these abstract mathematical ideas developed and became more widespread (via co-constructed shared intentionality; see Chapter 3), symbolic graphic conventions appeared to co-evolve (Figure 6-1; Figure 6-2, transition from Part B to C). As symbolic conventions became established (due to the co-construction of shared intentionality
among audiences), the P2: Abstract Symbol Prediction suggests that pictorial conventions would become controversial. According to the model, pictorial graphics would communicate the intended abstract concepts with more conceptual ambiguity (Concrete pictorial Prediction) than symbolic graphics would (P2: Abstract Symbol Prediction). The use of pictorial graphics would impede communication of the intended abstract concepts by consuming finite attentional resources that would otherwise be used in the construction of simulated prototypes of the intended abstract concepts (i.e., they would consume attentional resources by emulating pictorial properties; P3: Interference Prediction).

Instead, the communication of intended abstract concepts is more effectively afforded by symbolic graphics, which interfere less with the simulation of intended abstract concepts (P2: Abstract Symbol Convention). In simple terms, pictorial graphics that would have been acceptable in an earlier era (Figure 6-2, Part A) became controversial during Hilbert's era of the 19th and 20th centuries (Figure 6-2, Part C), ultimately resulting in a preference for (and more confidence in) a highly symbolic form that could afford the communication of intended abstract concepts.

As the fields of mathematics and logic continued to evolve, computer science emerged as a subfield of mathematics, and was closely linked with the field of electrical engineering (Carroll, 2013). Symbolic properties, employed to communicate abstract mathematical concepts, served as the foundation for modern programming languages (cf. Hoare, 1969; 1999) and were used by scholars in mathematics, engineering, and related fields (Figure 6-2, Part E). According to the model, the co-construction of the shared intentions required to use the symbolic conventions would have resulted in the appropriation of these symbols for the new field of computer science. It seems that it was only after computer use spread to fields outside mathematics and engineering (where the shared intentions would be different) that the symbolic properties of modern programming languages and interface displays became problematic. The field of human-computer interaction was born (cf. Carroll, 2012; Figure 6-2, Part F).

The designers of human-computer interfaces discovered that the 'visual' pictorial properties of graphical user interfaces (GUIs) made many computer tasks easier, more natural, and more intuitive for non-technical audiences (Carroll, 2013). According to the model developed in this dissertation, 'non-technical audiences prefer the pictorial properties of GUIs
because they use lower-level perceptual categories of vision that emerged 'naturally' within a physical environment (Concrete pictorial Prediction): Thus, the pictorial properties of the GUI seem 'more natural' or 'less learned.'

The success of the GUI probably influenced the design of 'visual' or 'pictorial' logical reasoning systems (e.g., Barwise & Etchemendy, 1992) and programming languages (e.g., Boshernitsan & Downes, 2004). However, these ‘visual’ or pictorial approaches have rarely been successful (as noted in Chapter 1; Blackwell, 2006), or are only successful in highly constrained situations (e.g., Green et al., 1991; see Whitley, 1997). Based on the model developed here (via the P3: Interference Prediction), the pictorial properties of visual programming languages will impede communication of the abstract concepts used in modern computer science.

6.9 Discussion in Relation to Prior Work

6.9.1 Response to Mumma

As discussed above, Mumma tried to demonstrate how Euclid's pictures did not introduce flaws into the Elements. I will now revisit their argument within the context of the model developed here and explore why, although their arguments are sound, pictures have still not been naturally selected in a way that would encourage widespread use. Recall how Mumma employed the term ‘co-exact property’:

...with respect to their co-exact properties the diagrams are all the same (Mumma, 2010, p. 223).

For two more items to be considered (e.g., perceptually processed as) ‘the same,’ they would require a higher-level category that they both fall under. Furthermore, if these items that fall under a common higher-level category could be differentiated, then the differentiations are the result of subcategories within the higher-level category. Applying the vernacular developed thus far, lower-level categories are more perceptual categories, whereas higher-level categories are more conceptual categories.

Let us now return to graphics. For two more pictures to be considered ‘the same,’ they would require a higher-level more conceptual category that they both fall under. Furthermore, if
these pictures that fall under a common higher-level conceptual category could be differentiated, then the differentiations are the result of lower-level (more perceptually specific) subcategories within the higher-level (more conceptual) categories. If Mumma’s co-exact properties are what is ‘the same’ about the diagrams, then co-exact properties are ‘the same’ as (composed of, or akin to) the conceptual categories that the pictures are considered to fall under. Applying the vernacular developed thus far, if the author of a graphic intended to communicate the conceptual category under which the graphic falls, then the relevant aspect of the graphic is its symbolic properties. In other words, co-exact properties are the same as, or are akin to, symbolic properties.

Let us now recall how Mumma employed the term ‘exact property’: “For any two magnitudes of the same type, one will be greater than another, or they will be equal. These relations comprise the exact properties of the diagram” (Mumma, 2009, p. 223). Because exact properties are what can be different about pictures of the same type, then exact properties appear to use more perceptually-specific categories within higher-level, more (abstract and) conceptually specific categories. Because pictorial properties are an author's intended emulation of a graphic, exact properties appear to be the same as, or are akin to, pictorial properties.

Thus, according to the model, Mumma is correct: it is possible for Euclid's pictures to form part of a valid logical argument. However, the model also reveals that it is the symbolic properties of Euclid’s pictures that are relevant to this argument. As demonstrated by the three predictions, pictures are not effective symbols: they consume resources required for the simulation of abstract concepts, and also introduce pictorial and perceptual specificity that obscure abstract concepts.

6.10 Chapter Conclusion

This chapter was a first attempt to demonstrate how the model developed here could be applied to explain phenomena in the domain of graphics. The P2: Abstract Symbol Prediction (sentences afford the communication of abstract concepts more effectively – or with more certainty – than pictures) and the P3: Interference Prediction (pictorial information interferes with symbolic information) were demonstrated using a site of artefact evolution: the artefact evolution of Euclid's *Elements*.
This demonstration involved tracing the natural selection of the *Elements* to a symbolic paradigm from the ‘primordial soup’ of pictures, in an effort to clarify whether and how symbolic information is more effective than pictorial information at affording the communication of abstract concepts. The model helped clarify the mechanisms and processes that drove this natural selection, thus demonstrating the model’s efficacy.

More specifically, the demonstration began by tracing the evolution of sentences from pictures, starting with the role of pictures in affording the communication of concrete concepts via ancient agricultural land-surveying practices. Symbolic graphics emerged over time, affording the communication of the more abstract mathematics and logic of Euclid's *Elements*. The *Elements* remained the exemplar for mathematical argumentation until the 19th century, when Hilbert and others identified a 'weakness' in the field of mathematics, and attributed this to Euclid's use of pictures to support key parts of the argument in the *Elements*. Mathematics was then re-invented upon a symbolic foundation in the 19th century. Key elements of the model developed here were used to show how this natural selection of sentences was driven by the need to communicate abstract concepts with more conceptual certainty.

The significance of this chapter relative to the goals of the dissertation is that it demonstrated and informally tested Explanative Prediction 1.

One of the ideas from the diagrammatic reasoning community reviewed in this chapter is the idea of ‘free rides’ (e.g., Shimojima, 1996), or how ‘diagrams reduce inferential load’ during reasoning. From this perspective, diagrams offer many advantages over text-sentences (e.g., Shimojima, 1996; Shimojima & Katagiri, 2008). However, if this is the case, the principles of artefact evolution raise a question: Why have diagrammatic proofs and visual programming languages never become widespread? The next chapter will respond to this question.
7  Demo #2: A Diagrammatic ‘Free Ride’ is not Truly Free: The Cost is Conceptual Ambiguity

7.1  Introduction

One of the ideas presented in the previous chapter was borrowed from the diagrammatic reasoning community: the idea that (pictorial) graphics have (via some set of metrics that is not entirely clear) certain advantages over symbolic graphics, and that these advantages relate to intuitive understanding and the ability to spontaneously draw new information from diagrams (cf. Burton & Coppin, 2012). One commonly cited advantage of diagrams is the ‘free ride’ phenomenon (Shimojima, 1996; Shimoima & Katagiri, 2008), which is related to the claim that diagrams reduce inferential load during reasoning (Barwise & Etchmendy, 1991). The ‘free ride’ phenomenon is often attributed to so-called ‘operational constraints’ that offer certain inferential steps ‘for free,’ for example:

Many, perhaps all, systems of diagrams have the function of letting the user [to] exploit spatial constraints on their graphical structure, and thus lightening the load of inferences. Consider a system of simple position diagrams, where letter symbols are arranged vertically to express a certain transitive relation holding between the symbolized objects. Figure 5.1 shows a sample diagram in this system, which expresses that the object A is lighter than the object B. Now, modify this diagram to express another piece of information, that the object C is lighter than the object A. We obtain the new position diagram shown in Figure 5.1.

This diagram is the result of expressing two pieces of information in the current system of position diagrams. Yet, it expresses a third piece of information, namely, the information that C is lighter than B. Furthermore, given the transitivity of the relation lighter, this additional piece of information is a logical consequence of the original two pieces of information. Thus, just by expressing the two premises in this system of position diagrams, the user obtains a diagram that expresses a particular logical consequence of them automatically. As aptly put by Barwise and Etchemendy (1991), the user “never need infer” this consequence from the premises, but “can simply read [it] off from the diagram as needed.”
Note that this “automaticity” of expression is largely due to a spatial constraint on the arrangement of letter symbols in a position diagrams: if a letter symbol x is placed above another letter symbol y, which is placed above still another symbol z, then the symbol x necessarily comes above the symbol z. The system of position diagrams is designed so as to exploit this spatial constraint for the “automatic” expression of certain logical consequences. The user can ride on this function of the system and significantly reduce his or her inferential task, replacing it with a reading-off task (p. 75).

Several researchers working within the field of diagramming have referred to the so-called ‘inferential advantage’ of diagrams, using terms such as ‘perceptual inference’ (Larkin & Simon, 1987), ‘non-deductive representation’ (Linday, 1988), and ‘emergent effects’ (Kulpa, 2003; Chandrasekaran et al., 2004). In general, the consensus in this community is that these advantages arise from the spatial constraints of graphics (Barwise & Etchemendy 1990; Shimojima, 1996; Stenning & Lemon, 2001).

However, if diagrams (and their spatial constraints) are truly advantageous, the principles of artefact evolution discussed here beg the question: Why have diagrammatic proofs and visual programming languages never become widespread in the communication of proofs (presumably through the natural selection of graphic representation types)? I will apply the model developed in this dissertation to explore this issue, and demonstrate that free rides are afforded by a diagram’s pictorial relations (among symbolic objects; ‘P5: Diagrammatic Prediction’), and are therefore not truly ‘free,’ given that pictorial relations impede the communication of symbolic relations (‘P3: Interference Prediction’). I will demonstrate that the ‘free ride’ actually has a cost: conceptual (and therefore, symbolic) ambiguity. This ambiguity is a disadvantage when the functional purpose is to communicate abstract concepts, which explains why diagrams were not naturally selected for proof communication (BSIP: Behavioural Symbolic Information Prediction from Chapter 5). However, this conceptual and symbolic ambiguity affords certain kinds of design and problem-solving, which explains why diagrams are seen as advantageous.

Before continuing, it is important to provide some context, specifically the discussion within the diagrammatic reasoning community about the potential limitations of diagrams. Researchers argue that although diagrams may offer many advantages, diagrams may also impede certain kinds of communication or related logical reasoning activities. Shimojima (1996;
2004) called this ‘over specificity’ – the “incapacity of the system to express certain sets of information without choosing to express another, non-consequential piece of information” (Shimojima, 2004, p. 18). By employing the P5: Diagrammatic Prediction, I will demonstrate how the pictorial relations of diagrams convey concrete relations that may interfere with the abstract relations intended by an author.

**Outline.** To investigate the ‘free ride’ phenomenon within the context of the model developed here, I will first demonstrate that operational constraints are akin to pictorial properties. More specifically, by applying the model’s C4: Diagrammatic Claim, I will demonstrate how the operational constraint of a diagram is a pictorial relation (among symbolized objects). Expanding on this, I will apply the P1: Concrete Pictorial and P5: Diagrammatic Predictions to show how the ‘free ride’ (or ‘inferential properties’) of a diagram is a product of the conceptual (and therefore symbolic) ambiguity of the diagram’s pictorial relation. In other words, the ‘ride’ is not truly ‘free’ – the cost of the ‘ride’ (the inference that is afforded by a pictorial relation) is that the represented relation is conceptually (and therefore symbolically) ambiguous (P5: Diagrammatic Prediction). Next, I will show how this conceptual (and symbolic) ambiguity affords certain kinds of logical system design and problem-solving, because each pictured relation can have multiple conceptual (and symbolic) meanings. This is basically another way to describe the free ride phenomenon; by applying the information theoretic and perceptual-cognitive aspects of the model, the true cost of a free ride emerges. Although this conceptual (and symbolic) ambiguity can assist in problem-solving and logical system design, it is a detriment to the communication of abstract concepts in math and logic. I will argue that the reason why diagrams have not been naturally selected for proof communication is that they are conceptually (and symbolically) ambiguous, but also that this conceptual ambiguity affords advantages during the design of logical structures (and how free rides are afforded by conceptual ambiguity).

### 7.2 Part A: Operational Constraints as Pictorial Properties

First, let us consider how operational constraints are akin to pictorial properties. According to Barwise and Etchmendy’s classic characterization of operational constraints:

> Diagrams are physical situations. They must be, since we can see them. As such, they obey their own set of constraints ... By choosing a representational scheme appropriately,
so that the constraints on the diagrams have a good match with the constraints on the
described situation, the diagram can generate a lot of information that the user never need
infer. Rather, the user can simply read off facts from the diagram as needed (Shimjoima

If operational constraints are the physical properties of a diagram (Barwise &
Etchemendy, 1991, referred to this as ‘physical situations’), relevant properties of a diagram
would include, for example, marks on a page, reflecting light that is emulated at a point of
observation. In essence, one operational constraint in this diagrammatic reasoning context is the
pictorial property that produces pictorial information (C1: Pictorial Information Claim).

However, equating operational constraints with pictorial properties (or information) is not
enough to explain the unique roles afforded by diagrams in reasoning and free rides. It is difficult
to imagine a logical reasoning diagram that does not include symbols, such as the labels used to
identify parts of a diagram. Labels A, B, and C of Figure 3.1 could symbolically represent the
lightness of various items, as in Shimojima’s (2008) example quoted above (where the relative
lightness of each item is represented via a pictorial relation, or what Shimojima called a ‘spatial
relation’). However, A, B, and C could just as easily symbolically represent different blocks in a
stack (see Figure 3.1). Simply, the labels used in these examples symbolically represent (often
abstract) objects, whereas the relations among the objects are represented pictorially. The
diagram is a pictorial relation among symbolized objects, corresponding to the C4:
Diagrammatic Claim. Let us now consider how pictorial relations afford free rides.

7.3 Part B: Pictorial Relations are Conceptually Ambiguous

The previous section described how the operational constraints of diagrams are akin to
pictorial relations (among symbolized objects). Let us now recall what pictorial relations of
diagrams are predicted to afford. According to the P5: Diagrammatic Prediction, the pictorial
relations of diagrams afford the communication of concrete relations with more certainty, and
abstract relations with less certainty (compared with symbolic relations). In other words,
pictorial relations are conceptually (and therefore symbolically) ambiguous (compared with
symbolic relations). Using the information theoretic terms introduced in Chapter 5, the concrete
pictorial relations of diagrams are more conceptually ambiguous because they have multiple
conceptual (and therefore symbolic) meanings. Let us now consider how this conceptual ambiguity can aid a logical problem-solver:

An author trying to compose (design) a sentence often considers a variety of possible conceptual relations among conceptual objects. A representation that simultaneously represents a range of possible conceptual objects could therefore aid the selection of a suitable conceptual relation from a range of possible conceptual relations. The next section explores this idea in more detail.

7.4 Part C: Conceptual Ambiguity affords Free Rides

This section focuses on how conceptually ambiguous relations offer advantages for logical system design and problem-solving, but disadvantages for the communication of abstract concepts of mathematics and logic. First, let us consider what design entails.

The vast body of research about design cognition (e.g., Cross, 2001) is beyond the scope of this dissertation; I will simply recruit a classic and widely used definition of design as a search of the space for possible solutions for the best or ‘satisficing’ solutions (Newell, 1969). Applying this definition to the design of logical arguments could mean that these ‘possible solutions’ are potential symbolic objects or relations that perform roles in the logical argument. What type of graphic represents a space for possible symbolic (abstract) relations? An ambiguous conceptual relation is a relation with multiple possible conceptual meanings or interpretations. Thus, because pictorial relations are predicted to be conceptually (and symbolically) ambiguous (compared with symbolic relations), they are therefore ideal for the design, or selection, of suitable symbolized abstract relations from the set of all possible symbolized abstract relations. Let us now apply this prediction to a real-world example of logical system design.

7.5 A Real World Design Example

To illustrate how conceptual ambiguity could afford (the early and intermediate phases of) logical system design, consider this example from Coppin (2010): use of graphic representation during software design. As shown in Figure 7-1, an information architect sketches outline drawings during the early design phases. As her ideas become increasingly refined, symbolic elements are introduced to represent abstract objects (databases, users, etc.). Next, the
picted relations among those symbolic objects are symbolized. Finally, symbolic relations emerge among symbolic objects. According to the C5: Sentential Claim, these symbolic relations are sentences.

**Figure 7-1: Design problem-solving in interaction design.**

Based on this example, it seems plausible that the design of an abstract concept, such as a mathematical proof or a computer program, could be a process of selecting a particular abstract concept from the set of possible abstract concepts. Because pictured relations afford the simulation of multiple possible abstract (and symbolic) concepts, pictorial relations of diagrams are probably advantageous, providing inferential steps 'for free.'

However, in the end, a 'free ride' is not actually free – it is a trade-off, because the price of a free ride is conceptual ambiguity. The next section explores this claim within the context of related research from the diagrammatic reasoning community about the limitations of diagrams.
7.6 Part D: Comparing the P3: Interference Prediction to Content Specificity

As noted in the Introduction, some theorists have suggested that diagrams engender ‘unwanted specificity’ – the “incapacity of the system to express certain sets of information without choosing to express another, non-consequential piece of information” (Shimojima, 2004: 18). Content specificity has been described as the reason why it is so difficult to express ‘abstract’ information via graphics (e.g., Stenning & Oberlander, 1995; Shimojima, 1996).

An example (directly quoted below) will help illustrate this idea of content specificity; Shimojima (1996, pp. 45-49) referred to a memory map in an individual named Harry (drawn from Hohauser, 1984):

…Recall that Harry has constructed the following memory map on the basis of his somewhat fragmentary memory about his home village:

![Figure 7-2: Harry’s Memory Map from Shimojima (1996).](image)

Harry now recalls that:

(θ1) The house B was somewhere between the house A and the house K

Before forgetting it again, he wants to record this information in his map by placing the B block between the A block and the K block. However, this requires him to put the B
block either between the A block and the road line 1, or between the road line 1 and the L block, or between the L block and the K block. He cannot decide which alternative to take because his memory does not tell him which of the following was true:

(θ2) The house B was between the house A and the road 1
(θ3) The house B was between the road 1 and the house L
(θ4) The house B was between the house L and the house K

Harry’s trouble comes from the content specificity of the system of his memory maps. Let $\Theta_m$ be the set of pieces of information presented in the map depicted in Figure 7.2. ([Shimojima] assume[s] that there are fixed rules of interpreting his memory maps.) The problem is that Harry’s memory maps cannot present the set of information $\Theta_m \cup \{θ1\}$ in isolation, without presenting either $θ2$ or $θ3$ or $θ4$ at the same time. In other words, any representation in Harry’s memory map system must be specific about the choice among $θ2$, $θ3$, and $θ4$ if it is to present the information $\Theta_m \cup \{θ1\}$.

… the over-specificity of Harry’s system can lead the representations to hold so-called “accidental features.” Although Harry actually knows better, let us suppose he went ahead and placed the B block between the A block and the K block, choosing a particular location, say, between the A block and the road line 1. The following map would result:

![Figure 7-3](image_url)

Figure 7-3: Another figure of Harry’s Memory Map from Shimojima (1996).
This map supports not only the state of affairs:

(σ1) The block B between the A block and the road line 1 but also:

(σ2) The block B is across the C block over the “river” band.

(σ3) The block B is closer to the bridge icon P than the L block is.

(σ4) The block A is closer to the B block than the L block is.

Note that the states of affairs σ2, σ3, and σ4 are all consequences of Harry’s arbitrary choice of putting the B block between the A block and the road line 1—they are accidental features of the map. Nevertheless, they have the following independent semantic values respectively:

(θ5) (θ6) (θ7)

The house B is across the house C over the river.

The house B is closer to the bridge P than the house L is. The house A is closer to the house B than the house L is.

This is why the above accidental features are potentially dangerous. In the course of his problem solving, Harry may forget that σ1, σ2, σ3, and σ4 are accidental features, and may “appeal to” one of them to base his future reasoning on the information indicated by it.

To consider this example in relation to the model, consider that Henry’s memory of the house was expressed linguistically, an abstraction from the particular concrete instances from which it was originally perceptually processed: “(θ1) The house B was somewhere between the house A and the house K.” (Shimojima, 1996, p. 46 [emphasis added]).
Table 2: Deconstructing the example of over specification from Shimojima (1996).

<table>
<thead>
<tr>
<th>Objects</th>
<th>&quot;the house B...&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;the house A and the House K...&quot;</td>
</tr>
</tbody>
</table>

The house only became over specified (using Shimojima’s terminology) when it was placed on the map as a pictorial (and therefore concrete) property. See Table 2: the relation “was somewhere between” cannot be expressed through a specific perceptual mode such as vision, sound, or touch – it is more amodal. Thus, according to the model developed here (adapted from Barsalou, 2003), the relation shown in Table 2 is an *abstract* relation. Representing this abstract relation via a pictorial relation would interfere with the communication of the intended abstract concept (P3: Interference Prediction).

Thus, the model developed here adds to Shimojima’s account. If an author’s intent is to provide information to guide an audience through a specific physical location, pictorial relations (e.g., of a map) would afford the desired action. However, if the intention is to communicate an abstract concept, such as “was somewhere between,” then a pictorial relation would interfere with the author’s intention. A symbolic relation would afford the author’s intended audience reaction more effectively (with greater certainty).

### 7.7 Conclusions: No Ride is Free: The Price is Conceptual Ambiguity

This chapter applied the model to show that a ‘free ride’ is a product of the conceptual ambiguity of a pictorially represented relation. The affordance of this conceptual ambiguity is multiple possible conceptual relations. This is a benefit in the field of design, but a detriment in the communication of proofs, which explains why diagrams are naturally selected in logical system design, but not in logical system communication.

First, I applied the model to distinguish diagrams from sentences, by conceptualizing diagrams as pictorial relations among symbolic objects, and sentences as symbolically represented relations among symbolically represented objects. This enables the prediction that diagrams afford the communication of concrete relations among symbolized conceptual objects,
whereas sentences convey concrete relations among concrete objects. The concrete relations of diagrams have ambiguous conceptual meanings; one advantage afforded by this ambiguity is that these relations can have multiple possible conceptual/symbolic meanings.

Thus, a free ride results from the conceptual ambiguity of a pictured relation; the affordances are multiple possible conceptual relations. This is a benefit in the field of design, but a detriment to the communication of proofs.

7.8 Epilogue for Chapter 7: Back to the *Elements*: A Free Ride Over a Long Time Scale

Now that the model’s claims and predictions have been applied to the cost of free rides, let us return to the natural selection of Euclid’s *Elements* and the axiomatic method from the ‘primordial soup’ of pictures, to update the findings of my previous research with collaborators.

Chapter 6 described a long process of artefact evolution, which began with pictorial graphics that afforded the communication of concrete structures for land surveying tasks, continued through the controversies of the 19th century (that engendered the reinvention of mathematics), and to the modern applications (and failures) of visual programming languages in contemporary human-computer interaction context, despite the prevailing view that ‘visual’ graphics are easier to use. In simplistic terms, this could be considered the ‘free ride’ phenomenon, stretched over a very long period of time.

7.8.1 Revisiting and Updating the Original Attention Theory

My previous research yielded three main findings:

**Coppin and Hockema (2009) Explanation 1: Why a diagram cannot usually constitute a convincing ‘holistic’ proof:**

Pictorial relations are more conceptually ambiguous than symbolic relations (P5: Diagrammatic Prediction). Furthermore, pictorial relations of diagrams interfere with simulation of P2: Abstract Symbolic objects, which is required to construct symbolic meaning via a graphic (P3: Interference Prediction).
Coppin and Hockema (2009) Explanation 2: *Why text is effective for proofs:*

Symbolic relations are more conceptually certain than pictorial relations (P6: Sentential Prediction). Furthermore, symbolic relations of sentences interfere less (compared with pictorial relations of diagrams) with simulation of P2: Abstract Symbolic objects, which is required to construct symbolic meaning via a graphic (P3: Interference Prediction).

Coppin and Hockema (2009) Explanation 3: *Why proofs are often sequential:*

Based on the P3: Interference Prediction, if the purpose of a graphic presentation is to communicate an abstract concept, pictorial information is predicted to interfere with the construction of the simulations required for the construction of symbolic meaning. For this reason, graphics in which pictorial objects and relations have little or no meaning would be naturally selected.

The model presented here supports these previous findings. In 2009, we suggested that focal attention must ‘walk’ through different parts of a conceptual visual-spatial structure due to a ‘narrower,’ but more intense focus (rather than attending to the whole structure at once).
8 Conclusions

This chapter summarizes the contributions of this work, discusses some limitations, and proposes directions for future research.

8.1 Summary and Limitations

In this dissertation, I developed a theoretical model for understanding graphic representations, distinguishing them in terms of their perceptual-cognitive affordances. This theory is a departure from previous research about representation: it does not treat representations as occurring ‘in the head’ (as in the cognitive sciences) or as occurring merely ‘in the media’ (as in the arts), but rather characterizes graphic representations as emerging from the interactions between the media and our biologically based and learned perceptual abilities. I explored various aspects of this process: how properties of graphic representations can be distinguished and understood in terms of biologically grounded and learned perceptual-cognitive capabilities; how these capabilities can be understood in terms of the dynamic environments within which they are naturally selected; and how graphic representations are naturally selected and evolve to perform functional roles as an interface between perceptual capability and a state of affairs, based on capabilities, contexts, and their purposes.

The theoretical model developed here provides insight into why some so-called pictorial properties of graphics might seem to inform actions more ‘naturally’ – with less effort, less learning, and less training. Additionally, the model provides insight into why some pictorial properties fail to afford certain kinds of actions (such as the communication of abstract concepts) that are more easily afforded by the so-called symbolic properties of graphics that might seem akin to spoken or written languages. Finally, the model provides insight into why some graphics, such as diagrams, seem to be hybrids composed of pictures and symbols (such as text-labels).

This dissertation is one step in a larger process that is intended to contribute to a science of visual information design. The model is a deductive and analytic tool that can help clarify what I uncovered through a previous inductive step: from 2007–2011, I conducted a series of concatenated studies (Stebbins, 2006; see e.g., Coppin & Hockema, 2008; Coppin, 2008b; Coppin, 2009; Coppin & Hockema, 2009; Coppin, Burton, & Hockema, 2010; Coppin, Kennedy, & Hockema, 2010; Coppin, 2011). This research sparked preliminary ideas, which were revisited
and updated by the model (primarily in Chapters 7–8). It also helped me develop the question posed at the beginning of Part II: “Why are (Geometric) Proofs ‘Non-Visual’?”

To constrain this project to a dissertation scale, I did not take the additional step of testing the model’s predictions experimentally; this will be a focus of future work. Additionally, to meet my objective of developing a scientific model that could incorporate the proto-theories of designers as well as theories about perception and cognition, I only drew upon research that could inform the development of my model (and that was compatible with affordance-based ideas of contemporary design). In future research I will focus on other relevant ideas from cognitive science, educational psychology, communication theory (especially the Toronto School of Communication), semiotics, and critical theory.

Despite these limitations, the model for graphic representation contributes to a scientific understanding of graphics in a number of ways. The model’s account of the constructed nature of perceptual capabilities used to perceive graphics is highly compatible with ideas from art theory, critical theory, and semiotics that historically have not been considered compatible with ideas from cognitive science and perceptual psychology. It offers a novel account of graphics and affordance-based ideas related to design by conceptualizing external graphic representations as subsets of an environment that is configured to cause intended percepts. It demonstrates how if graphics emerge at this intersection of perceptual capability and intentionally configured environment, then they can be understood in terms of how they use capabilities that were naturally selected within dynamic, and social, environments. In this way, new media (such as the information visualization-visual analytics used in human-computer interfaces and the web) can share common properties with old media (such as comic strips, outline drawings, books, and cave paintings). These common properties are a function of perceptual-cognitive capabilities used by both old and new media.

The model distinguishes between pictorial, diagrammatic, and sentential graphics in terms of how each type of graphic pictorially or symbolically represents objects and relations among objects. Two fundamental perceptual-cognitive properties of graphics were conceptualized for this purpose: pictorial information (Claim 1) and symbolic information (Claim 2). The model distinguishes between pictorial and symbolic information based on two interrelated aspects of an individual’s homeostatic perception-reaction loop, a concept that was
synthesized and modelled from the literature. ‘Emulation’ refers to the aspect of perception-reaction that ‘picks up’ or detects produced or reflected light that impinges upon sensory receptors. ‘Simulation’ refers to the aspect of perception-reaction that processes, filters, or interprets that which is emulated, and thereby enables an organism to anticipate changes and variations that are relatively more distal, in terms of space and time. Conceptualizing pictorial and symbolic information also required developing a way to conceptualize information in a graphic, in contrast to information in an environment: A (graphic) representation was conceptualized as an environment that has been intentionally configured to produce (visual) information that causes an intended percept.

The model predicts that pictorial information affords the communication of concrete concepts (more effectively than symbolic properties; Pred. 1); that symbolic information affords the communication of abstract concepts (more effectively than pictorial properties; Pred. 2); and that pictorial information interferes with the communication of abstract concepts via symbolic information (Pred. 3).

By focusing on Claims 1–2 and Predictions 1–3, I developed a way to distinguish between pictures, diagrams, and text in terms of how objects or relations are pictorially or symbolically represented. According to the model, diagrams are pictorially represented relations among symbolically represented objects (Claim 4); sentences are symbolically represented relations among symbolically represented objects (Claim 5); and pictures are pictorially represented relations among pictorially represented objects (Claim 3).

By integrating these claims and predictions, it was possible to predict that diagrams afford the communication of concrete relations among abstract objects (Pred. 5); that sentences afford the communication of abstract relations among abstract objects (Pred. 6); and that composite pictures afford the communication of concrete relations among concrete objects (Pred. 4). I was able to demonstrate how predicted affordances correspond to functional purposes by examining the evolution of graphics in a situation where we can be confident that the types of graphics observed have been naturally selected to effective states relative to functional purposes.

When the objective is to communicate an abstract concept, the model's explanatory prediction is that symbolic properties will be more effective (will communicate abstract concepts
with greater certainty) than pictorial properties (Concrete Symbolic Prediction). Based on the principles of artefact evolution, symbolic properties are more likely than pictorial properties to be naturally selected when the functional purpose is to communicate abstract concepts.

When the objective is to communicate a concrete concept, the model’s explanatory prediction is that pictorial properties will be more effective (will communicate concrete concepts with greater certainty) than symbolic properties (Concrete Symbolic Prediction). Based on the principles of artefact evolution, pictorial properties are more likely than symbolic properties to be naturally selected.

When the objective is to design a logical structure (such as a symbolically represented proof), the model’s explanatory prediction is that pictorial properties will be predominant in the early design phases, diagrammatic properties will be predominant in the intermediate phases, and symbolic properties will be predominant in the final phases.

8.2 Practical Implications for the Design of Graphic Displays

One potentially important implication of the model is that it predicts how pictorial properties can interfere with the communication of abstract concepts (P3-Interference Prediction). If the P3-Interference Prediction is well-founded, then the creation of a ‘visual’ or pictorial programming language may be less successful than symbolic alternatives. This finding could have a significant impact in visual programming language design communities (e.g., the IEEE Visual Language and Human Centered Computing Conference series) and diagrammatic reasoning communities (e.g., the Diagrams conference series), which are trying to develop more effective programming tools by employing pictorial graphics.

The overall practical implication of the model is that it suggests how a designer should determine when to employ pictorial or symbolic properties to afford the communication of concrete or abstract concepts. More specifically:

Implication 1. The first implication is that the model can help a designer decide when to communicate an intended object or relation through pictorial (as conceptualized in Claim 1) or symbolic (as conceptualized in Claim 2) means:
• **Implication 1a.** If an author intends to communicate a concrete structure, Predictions 1–3 anticipate that the intended concrete structure will be more effectively conveyed pictorially instead of symbolically.

• **Implication 1b.** If an author intends to communicate an abstract concept, Predictions 1–3 anticipate that the author’s intended abstract concept will be more effectively conveyed symbolically instead of pictorially.

**Implication 2.** The second implication is that the model can help a designer understand under what conditions relations among objects should be represented pictorially or symbolically (to produce composite pictures, diagrams, or sentences):

• **Implication 2a:** If an intended object is concrete, then the P1: Pictorial Prediction anticipates that the object will be more effectively represented pictorially.

• **Implication 2b:** Alternatively, if an object is abstract, then the P2: Symbolic Prediction anticipates that the object will be more effectively represented symbolically.

• **Implication 2c:** If a relation among objects is abstract, then the relation would be more effectively represented symbolically.

• **Implication 2d:** Alternatively, if a relation is concrete, then the relation would be more effectively represented pictorially.

Implications 3a–3c integrate Implications 2a–2d above:

• **Implication 3a:** Integrating the above, the P5: Diagrammatic Prediction anticipates that concrete relations among abstract objects will be more effectively communicated C4: diagrammatically.
• **Implication 3b:** Integrating the above, the P6: Sentential Prediction anticipates that intended abstract relations among abstract objects will be more effectively communicated C5: sententially.

• **Implication 3c:** Finally, the P4: Composite Picture Prediction anticipates that concrete relations among concrete objects will be more effective if represented via a C3: Composite Picture.

### 8.3 Broader Implications for Design Theory and Future Work

As discussed in the Introduction, the applied art and design field lacks foundational theories (e.g., Norman, 2010). The model developed here may be extended to develop more foundational theories for visual interface design. The next subsections will discuss how the homeostatic perception-reaction loop in the model, which can clarify affordances of graphics at the scale of an individual perceiver, can also be extended to clarify the role of graphically afforded homeostatic perception-reaction loops on a cultural scale. In addition to discussing the broader implication of this model, I will also briefly foreshadow how such an approach could help a designer understand when a designed artifact is effective, accessible, or inclusively designed (for diverse audiences) by clarifying how given artifacts facilitate shared perception-reaction loops among minds.₃⁹

Let us first recall the example from Chapter 3, which described how sharing homeostatic perception-reaction resources among individuals can contribute to survival and prosperity within environments of dynamic change and variation. In the telerobotics and ship navigation examples from Chapter 4, experiences of an Organism A (e.g., a previous explorer) were used by another Organism B (e.g., a crewmember of the HMS Bounty) after the results of prior expeditions had been graphically represented in charts and maps and disseminated throughout the British Navy. This example showed how the crew of the HMS Bounty, and the British Navy more widely,

₃⁹ This future work will explore what graphics (and designed artifacts more generally) afford, reducing the need to consider the role of intentionality (as was done in Chapter 3) when distinguishing graphically represented items from non-represented items.
were beneficiaries of previous perception-reactions that had been externalized and shared via graphics of charts and maps. For example, pictorial properties of an outline drawing depicting the concrete structure of coastal landforms could help a navigator chart a path to avoid obstacles that could sink or ground a vessel. This ‘shared perception-reaction loop’ between earlier explorers and later ship crews provided an advantage for the users of these graphic representations (relative to competitors who were less able to share homeostatic resources). Let us next consider how symbolic properties, and sentences in particular, could be understood as vehicles that convey abstract concepts among minds through these homeostatic loops.

Sentences, where both objects and relations are symbolized (C3: Sentential Claim) emerged as the preferred type of graphic for communicating concepts in mathematics and logic (see the discussion of Euclid’s *Elements* in Chapters 6–7). To consider why sentences may be a more effective vehicle for communicating theories among minds (compared with C1: Composite Pictures or C2: Diagrams), let us consider how theories have been used as a metaphor for simulations of the mind:

… mental representations [simulations] are more like theories: pieces of knowledge that can support many inferences in many different situations. For instance, Newton's theory of motion makes predictions about infinitely many different configurations of objects and can be used to reason both forward in time and from the consequences of an interaction to the initial state. The generative approach posits that mental representations [simulations] are more like theories in this way: they capture more general descriptions of how the world works—hence, the programs of the mind are models of the world that can be used to make many inferences. (Goodman et al., 2012)

Future work will extend the model developed here to show how this excerpt from Goodman (2012) is more than just a metaphor. Common properties between internal simulations (of individual minds) and theories (that are shared among minds) could emerge because of a common state-of-affairs shared by both individuals and groups: dynamic environmental change and variation. From this perspective, simulation of possible change and variation enables planning and prediction that engenders survival and prosperity in the face of dynamic change and variation at the scale of an individual organism, and theories enable planning and prediction that
engenders survival and prosperity in the face of dynamic change and variation at a *cultural scale*. Consider Popper’s description of a scientific theory:

> Scientific theories are universal statements. Like all linguistic representations they are systems of signs or symbols. Theories are nets cast to catch what we call ‘the world’; to rationalize, to explain and to master it. We endeavor to make the mesh even finer and finer. (1980, p. 59)

What is the vehicle that enables the sharing of theories across space and time? According to the model, *sentences*, and their symbolic relations among symbolic objects, afford the sharing of theories that are conceptual in nature (composed of *conceptual relations* among *conceptual objects*). Future work will show how the natural selection of sentences formed a ‘primordial soup of pictures,’ which acted as the basis – or was at least related to – the natural selection of the axiomatic method, which is the foundation for predictive theories in the sciences, social sciences, and engineering, as well as the conventions used in contemporary mathematics, computer science, and programming languages. Thus, the model can be extended to show how the theoretical structures that are naturally selected on a cultural level (and the graphics used to share them) are based on mechanisms and processes that can be modelled on an individual level. This may yield a more unified theory for design.

Let us now consider the practical implications of these finding within emerging areas of design, by considering what happens when institutions (and cultures) fail to facilitate the sharing of perception-reaction resources among individual minds.

### 8.3.1 Implications for Inclusive Design

Interfaces, such as computer screens, are critical to help individuals access resources to survive and thrive in everyday life. However, not everyone can access these resources with equal ease. Audiences for interfaces are diverse, and may differ in terms of physical needs (e.g., those related to age or impairment); culture or language; and training and expertise. Currently, approximately 20% of digital media consumers do not access graphics using vision: these vision-impaired users may use digital media via text-to-speech audio or Braille. The Web Content

---

40 For an overview of inclusive design see Clarkson (2003).
Accessibility Guidelines (http://www.w3.org/WAI/intro/wcag.php) were developed to make digital media accessible to this population. According to these guidelines, a visual graphic is considered ‘accessible’ if it has been ‘translated’ into a text description that can be read to an audience via screen reader technology. This approach to accessibility seems to ignore the objectives of those who create the visualizations: if text descriptions can adequately convey a visualizer’s intent, why create a visualization in the first place?

Future work will apply the model to clarify what is lost or gained when a visual is described using text, how what is lost may be delivered through another sensory mode (such as sound), and how individual differences shape perception of these media.

This model could lead to an alternative approach to presenting visualizations. For example, the model’s principles could inform the inclusive design of non-visual interfaces to the web. A screen-readable text description (which is C2-symbolic and C5-sentential, according to the model) of a visualization (which is P1-pictorial, according to the model) is predicted to convey the s3-concrete structures intended by the visualization’s author less effectively, but would instead convey s4-abstract interpretations introduced by the author of the text description.

Let us now consider how the model suggests that an auditory signal could more closely convey what might have been intended by the author of a visualization. In particular, the model’s P1-Concrete Pictorial Prediction suggests that C1-pictorial graphics (of a visualization) afford the communication of s3-concrete concepts with more specificity than C2-symbolic graphics. The strategy shown in Figure 5-4 suggests how C1-pictorial graphics are not the only type of perceptually certain (recall the discussion of perceptual certainty in Chapter 5) external representation. A tactile or auditory signal can also be perceptually certain (Figure 5-4). In this way, the model can be integrated with research about sonification (e.g., Herman et al., 2011) to produce more perceptually certain external representations that more accurately convey the s3-concrete structures intended via the P1-pictorial properties employed by the author of a visualization. Emerging preliminary work with my students and research assistants in this area is yielding promising results, for example by applying the model’s principles to produce more perceptually specific ‘song maps’ (sonifications) that are intended to convey the s3-concrete structures of a physical road or path (Figure 8.1; Crosskey, Campbell-Smith & Kwok, 2013; Coppin and Li, under development).
Figure 8-1: Sonifications of these roadmaps were produced by my students (Crosskey, Campbell-Smith & Kwok, 2013)

Together, the results will extend and apply the theoretical principles underlying the model, helping to develop display techniques that use non-visual sensory modes and thereby a more inclusive approach to presenting this kind of information. This future work will serve as an opportunity to extend and test the model by applying it to non-visual sensory modes such as sound, and thereby serve the needs of inclusive designers.

8.3.2 Broad Applications for the Model

To this point, my initial empirical work has helped develop a theoretical foundation and framework for graphic representation and cognition. My next step will be to use this framework as a ‘prism’ for additional qualitative and quantitative empirical research, in combination with selected applied graphic representation and/or system development.

Although my research agenda is strictly focused on graphic representation and cognition, my work is relevant to many other observation sites and applications, so I will be able to explore numerous research environments without deviating from my current focus (see Figure 8-2). Graphic representations are found in almost every imaginable human endeavor, and each example serves as a ‘sonar ping’ to expose phenomena that can be explored through my framework. I plan to exploit this fact to expand my research program through a combination of
applied graphic representation development (such as information design applied to real-world problems); cognitive ethnographies (to clarify the functional roles played by graphic representation in naturalistic settings); and selected experiments (to test theories developed through qualitative observations).

Figure 8-2: Potential research areas.

**Final words.** Contrary to the widely held view that pictorial graphics aid usability, problem-solving, and communication, the model developed in this dissertation reveals that pictorial graphics interfere with the communication of abstract concepts, but facilitate the communication of concrete concepts, and aid in the design of abstract symbolically represented structures. Additionally, the model shows how theories of perception and cognition can help synthesize competing theories of graphics. The overall goal of this work has been to develop a more general or unified theory that could become the basis for a science of representation – one that is synergistic with (and that can be operationalized within) current design practice, as well as related computer science and cognitive science fields involved in contemporary interface design.
9 Epilogue: Dual-Coding Theory in Relation to the P3: Interference Prediction

9.1 Introduction

The three predictions in this model (particularly P3: Interference Prediction) are related in some ways to Paivio’s (1991) dual-coding theory (DCT). Although the domain of multimedia learning (and educational psychology more broadly) is beyond the scope of this dissertation, a brief discussion is helpful here. This will be followed by an exploration of a possible case study: a site of artifact evolution where functional purposes are to communicate concrete structures. According to DCT:

…the human mind operates with two distinct classes of mental representation (or “codes”), verbal representations and mental images, and that human memory thus comprises two functionally independent (although interacting) systems or stores, verbal memory and image memory. Imagery potentiates recall of verbal material because when a word evokes an associated image (either spontaneously, or through deliberate effort) two separate but linked memory traces are laid down, one in each of the memory stores. Obviously the chances that a memory will be retained and retrieved are much greater if it is stored in two distinct functional locations rather than in just one. (Thomas, 2013)

Educational psychologists have applied dual-coding theory to the problem of so-called ‘multimedia learning,’ arguing that “Multimedia learning occurs when students build (internal) mental ‘representations’ from words and pictures that are presented to them” (e.g., printed text and illustrations or narration and animation; Mayer, 2003, p. 125). Multimedia-learning researchers developed a Multiple Representation Principal (MRP), arguing that it is “better to present an explanation in words and pictures than solely in words” (Mayer & Moreno, 2005). To validate the MRP, Mayer and Anderson (1991, 1992) tested students who listened to instructions for using a bicycle tire pump while viewing (or not viewing) a corresponding animation. They found that multiple representations generated twice as many useful solutions to subsequent problem-solving transfer questions compared with students who listened to the same instructions without viewing the animation. Similarly, students who read text along with captioned illustrations generated about 65% more useful solutions on a subsequent problem-solving transfer test compared with students who only read text (Mayer, 1989; Mayer & Gallini, 1990).
Do the claims of multimedia learning researchers raise doubts about the P3: Interference Prediction? According to the MRP, pictures and verbal descriptions are more effective than verbal descriptions alone; this section will show how the model developed here is consistent with the findings of multimedia learning researchers. It will go on to show how this model can actually contribute to multimedia research by providing a way to describe pictorial and symbolic affordances by distinguishing between concrete structures and abstract concepts. Multimedia researchers do not appear to make such a distinction. The implication is that MRP is a claim about the communication of concrete structures (for example, multimedia presentations about a break system and pump were the focus of the cited experiments); it may not apply to the communication of abstract concepts (as anticipated by Prediction 3). The concrete versus pictorial distinction offered by the model developed here may help clarify why other researchers (e.g., Byrne, Catrambone, & Stasko, 1999) have been unable to apply the findings from multimedia learning research to improve computer science educational materials focused on algorithms.

9.2 Dual-Coding Theory and Multimedia Learning: Sketch of Key Ideas

According to DCT, visual and verbal information are processed differently and along distinct channels, where they are ‘encoded’ into memory as separate ‘internal representations’ (referred to as ‘codes’ by DCT researchers). These codes are based on what was processed via each channel and organize that which was processed in ways that can be acted upon, stored, and retrieved for later use. For example, ‘dog’ could be stored as both a (spoken or written) word ‘dog’ and as a ‘mental image’ of a dog. According to DCT, a learner can retrieve the word individually, the image individually, or both simultaneously. One aspect of DCT that is of interest to multimedia learning researchers is its hypothesis that coding a stimulus two different ways increases the chance of remembering that item compared to if the stimulus was only coded one way; this is another way of describing the MRP.

9.3 Comparing the MRP to Predictions 1–3

This section compares the MRP to this model’s predictions. First, it is important to determine whether the materials tested by Mayer and Anderson (1991, 1992) were conveying abstract concepts, concrete structures, or both. They did not make such a distinction, but based
on the discussion of abstraction versus concreteness in Chapter 5, their tests appeared to involve concrete structures (recall the bicycle pump). However, descriptions of the pump’s processes would seem to require abstract concepts (such as the concept of air pressure or force). Therefore, the model developed here would recommend a hybrid presentation, composed of pictures (to convey the physical concrete structures of an air pump; P1: Concrete Pictorial Prediction), and sentences (to convey abstract concepts such as air pressure or force; P2: Abstract Symbolic Prediction). This allows it to synergistically predict Mayer’s results, but with a twist: it also suggests that Mayer’s claims are most applicable to presentations about concrete structures. However, according to P3: Interference Prediction, pictures will interfere when an educator’s intention is to communicate an abstraction alone (such as an algorithm’s principles). To date, no tests of MRP have been conducted that could be interpreted as challenging P3, but the unsuccessful extension of the MRP to the problem of algorithm visualization in computer science education (Stasko, Badre, & Lewis, 1993) suggests its limitations when the purpose is to communicate an abstraction.

### 9.4 MRP Frustration in Computer Science Algorithm Education

Inspired by the successes of multimedia learning, computer science educators Stasko, Badre, and Lewis (1993) attempted to apply the MRP within a new domain: algorithm instruction. They hypothesized that ‘visualizations’ or animations (‘more pictorial’ graphics) of algorithms would engender positive learning outcomes. In particular, they “anticipated that the animation would improve procedural understanding” (Kehoe, Stasko, & Taylor, 1999, p.1). However, they were disappointed to find that “animations did not perform any better than the control group on tests of their procedural knowledge” (Kehoe, Stasko, & Taylor, 1999, p.1).

They provided an explanation for their frustrating results, attributing the lack of a performance advantage via animation to a property shared by most visualizations: they generally represent an expert's understanding of the algorithm, not a novice's:

For a student to benefit from the animation, the student must understand both [the] mapping [from the algorithm to the graphics] and the underlying algorithm on which it is based.... Students just learning about an algorithm do not have a foundation of understanding upon which to construct the visualization mapping. (Kehoe, Stasko, & Taylor, 1999, p.1)
Later, Byrne, Catrambone, and Stasko (1999) went on to explore a more qualitative paradigm, to clarify how learners use of animations (instead of testing the effectiveness of animations). They were surprised to learn that:

one way animations may aid learning of procedural knowledge is by encouraging learners to predict the algorithm's behaviour. However, such a learning improvement was also found when learners made predictions of an algorithm's behaviour from static diagrams. This suggests that prediction, rather than animation per se, may have been the key factor in aiding learning in the present studies. (Byrne, Catrambone, & Stasko, 1999, p.253)

How can the model developed here explain Stasko et al.’s less-than-successful attempt to extend Mayer et al.’s work to algorithm visualization? According to this chapter’s interpretation and adaptation of Barsalou (e.g., Barsalou, 2003), algorithms are in the ‘abstract’ domain. According to Prediction 2, pictorial graphics are anticipated to be less effective than symbolic graphics when the functional purpose is to communicate an abstraction, partly because multiple perceptual (i.e., pictorial) structures (Figure 5-7, left) can fall within a given abstract conceptual category (Figure 5-7, right): In other words, pictorial graphics communicate an intended abstract concept with less certainty.

This model also can help explain why Byrne, Catrambone, and Stasko (1999) found that “one way animations may aid learning of procedural knowledge is by encouraging learners to predict the algorithm's behaviour”. Because a symbolically represented algorithm is intended to engender the construction of a more amodal audience simulation (Figure 5-9, right), compared with a more modal simulation or emulation of a concrete structure (Figure 5-9, left), many possible pictorial structures could fall within the symbolically represented abstract categories. In other words, a pictorial graphic (or set of graphics) intended to illustrate an algorithm’s behaviour, even if deemed accurate by a panel of experts, is only one concrete example among a range of possible concrete examples. In other words, a pictorial illustration of a symbolically represented abstract concept results in a concrete interpretation of the author’s intended abstract concept. This has three implications that future research could explore:
1. The model developed here anticipates how each student may produce a different (more modal) concrete interpretation (Figure 5-5, left and Figure 5-9, left) of a (more amodal) algorithm (Figure 5-9, right).

2. Producing *multiple concrete examples* (Figure 5-5, left and Figure 5-9, left) should be more effective (for a given student) than producing a single example, because what is *relevant* about an (amodal) abstract concept (Figure 5-8, right and Figure 5-9, right) is its ability to be applied to many concrete (more modal) examples (Figure 5-5, left and Figure 5-9, left). Thus, the production of multiple concrete *interpretations* (Figure 5-5, left and Figure 5-9, right) of the abstract algorithm (Figure 5-8, right and Figure 5-9, right) should produce a deeper understanding of a given abstraction (Figure 5-8, right and Figure 5-9, right). The abstraction (Figure 5-8 right and Figure 5-9, right) is intended for application across a variety of phenomenon (Figure 5-5, left and Figure 5-9, left).

3. A given concrete pictorial representation (Figure 5-5, Perceptual Category #1) could impede a student’s ability to grasp (or simulate) the relevant (more amodal) properties of an author’s intended abstraction (Figure 5-8, right), by drawing the student’s attention to a particular concrete *instantiation* (or *example*) rather than the intended abstraction that could be applied to explain *many possible concrete instantiations*. In other words, a pictorial representation of an abstraction may consume finite simulation resources that would otherwise be recruited to simulate the author’s intended abstraction.

**9.5 Conclusion: The Model is Synergistic with Dual Coding Theory and Can Also Contribute to It**

The model’s predictions are synergistic with the findings from Mayer and Anderson (1991, 1992), but only when a distinction is made between the communications of concrete structures versus abstract concepts. In their study, the concrete structures and properties of an air pump were afforded more effectively when pictorial graphics accompanied sentence-based descriptions.
10 References


Clarkson, J. (2003). Inclusive design: Design for the whole population. Location of Publisher?:Springer Verlag.


Coppin, P.W. (2009, February 9). Using dyslexia to explore the cognitive characteristics of illustrations and text; using illustrations and text to explore the cognitive characteristics of dyslexia. In proceedings of iConference 2009, University of North Carolina at Chapel Hill, Chapel Hill, NC.


Ware, C. (2010). *Visual thinking: For design*. Morgan Kaufmann. Location of Publisher.


Appendices

Appendix A: Proposition 35 of Euclid’s *Elements*

Proposition 35 of Euclid’s *Elements* is that parallelograms that are on the same base and in the same parallels equal one another. The description of the following example is adapted from Coppin and Hockema (2009). Euclid’s proof proceeds as follows:

![Diagram](image.png)

**Figure Appendix A-1: Diagram used in proof of Proposition 35.**

Proof.

i. Let ABCD, EBCF be parallelograms on the same base BC and in the same parallels AF, BC.

ii. Since ABCD is parallelogram, AD equals BC (Proposition 34). Similarly, EF equals BC.

iii. Thus, AD equals EF. (Common Notion 1)

iv. Equals added to equals are equal, so AE equals DF. (Common Notion 2)

v. Again, since ABCD is a parallelogram, AB equals DC (Proposition 34) and angle EAB equals angle FDC (Proposition 29).

vi. By side angle side congruence, triangle EAB equals triangle FDC (Proposition 4). Subtracting triangle EDG from both, we have that the trapezium ABGD equals the trapezium EGCF (Common Notion 3).

vii. Adding triangle GBC to both, we have that ABCD equals EBCF (Common Notion 2);

Note how the text references aspects of the visual-spatial concept sequentially. For example, the first line introduces the symbol “ABCD”, which, in the presence of the figure, directs attention sequentially through the vertices A-B-C-D and then to the parallelogram as a whole, separable from the rest of the figure. However, the figure is not necessary for this step, for ABCD could also serve the role of a simple “word” (term) in the logical proof without actually referring to the geometric figure at all. Further, for everything specified early on in the proof—
symbols and relationships—each line can be derived from the previous without making reference to the figure at all. Although the figure can still play a helpful illustrative role, it does not play a role in sanctioning particular inference steps. Indeed, as Mumma (2009) describes, this is the case all the way up to step iv.41

However, things get more complicated in steps iv through vii, in part, because the author of the proof (Euclid via translation) seems to be making the assumption that we will be using the figure to provide interpretation for the text statements.

41 Step iv contains a so-called “flaw” in that in order to determine what the “equals” are that have been added to AD and EF, one needs to realize that DE is a common segment shared by both AE and DF. The step relies on an understanding that DE has been “added” to both AD and EF and it is reflexively self-equal. However, while it is true, it was nowhere stated in steps i-iii that DE is a shared segment, i.e., that D lies on segment AE and that E lies on segment DF. In the above proof, this is knowledge that must be gleaned from the figure.