CONTINUITY IN CATEGORIZATION
AND THEORETICAL IMPLICATIONS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Richard A. C. Dale
August 2006

CONTINUITY IN CATEGORIZATION
AND THEORETICAL IMPLICATIONS

Richard A. C. Dale, Ph.D.
Cornell University 2006

Traditional theories of cognition assume that motor action is executed in an all-or-none fashion, and has little importance for understanding cognitive representation and processing. A series of experiments and simulations presented here challenges this assumption. A relatively higher-order cognitive process, categorization, is shown to have graded effects that are reflected in manual motor output, measured through streaming x-y coordinates from mouse trajectories. Two simulations show that these effects are likely generated from a system in which cognition and action interact fluidly. Finally, theoretical implications of these experiments are drawn out. Symbolic dynamics is introduced, a potential means for reconciling both traditional and continuous accounts of cognition. A broad philosophical discussion follows, in which an integrative and pluralistic approach to cognition is proposed and briefly discussed.

BIOGRAPHICAL SKETCH

Richard A. C. Dale was born in Toronto, Canada in 1977. He attended the University of Toronto and obtained a bachelors degree in linguistics. After some work at Southern Illinois University at Carbondale, where he pursued connectionist psycholinguistics research in the Department of Psychology, he moved to the Department of Psychology at Cornell University, where he will with some luck complete his graduate studies.

ACKNOWLEDGMENTS

I am eternally grateful to Morten Christiansen and Michael Spivey for giving me both wonderful encouragement, and unrelenting criticism, in guiding my continual development as a psychologist. Thanks also to the great community of professors and graduate students (and undergraduate researchers) in Cornell Psychology for providing a wonderful environment in which to undergo this development. I would like to thank Barbara Finlay, and the opportunity to work on this dissertation under the auspices of a Paller-Dallenbach Fellowship in my final year of graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

Introduction: Body and Mind

Broader Relevance

This dissertation addresses two broad concerns in cognitive science. The first is pervasive in debate, the second not so much at all. The first is a key concern about how cognitive processes unfold in time. This mostly centers on the dispute about the relative discreteness, or continuity, of cognitive states and processing. One traditional line of thought sees the cognitive system, in achieving some information-processing goal, transition through discrete representational states, much like a digital computer. Another, growing out of a long but more recently relevant history, sees the cognitive system smoothly move through probabilistic representational states, more like the flow of a dynamical system. Such a dispute might be described as pursuing the *matters of fact* about cognitive representation and process (as Hume, 1748/1967, famously noted about certain classes of knowledge; or, recently, as used by Quine, 1960). In other words, the debate seeks to definitively state the nature of the cognitive system by figuring out how it *actually, in reality, functions*. The ultimate description resulting from this debate is hoped to be exactly that: the ultimate, end-of-discussion, characterization of cognitive processing.

The second concern is arguably more broad, and seeks to compare these descriptive schemes in another, wider descriptive framework. The concern is thus metatheoretical. By employing some novel concepts from dynamical systems mathematics, these descriptive schemes may in fact end up being less competing than the first debate recommends. Instead, both discrete and continuous descriptions may serve important functions depending on the cognitive phenomena under investigation. Such a *pragmatist* approach to theoretical descriptions relies on the varying *units of analysis* available in cognitive representation and process (as Skinner, 1938, describes in uncovering functional units of behavior, or Dewey & Bentley, 1949, on designating appropriate units of action; see Palmer, 2003, for a review). In other words, depending on the units of "mind/brain" selected for study (for example in language: words, sentences, conversation, etc.), discrete descriptions or continuous descriptions may serve to better predict and explain systematic relationships among the behavioral variables under study.

In what follows, I trace a selective history of philosophical and scientific investigations of cognition, and how strong historical trends have led to the first debate. This selective survey lays out the fundamental questions about the matters of fact of cognition, the answers to which seek to discover the real nature of cognitive representation and process. The first few sections of the following discussion set the context for the empirical and computational work presented in subsequent chapters. Following this, I offer some brief preliminary reflections on the second broad concern. How might conceptual advances mitigate or change the nature of this debate? This sets the stage for later theoretical chapters, in which such a discussion is engaged directly.

To begin this historical survey, I consider the relationship between the body and the mind. The strongest historical trend in both the philosophy and science of the mind is to consider these two entities as quite distinct: Thinking is a species different from the body, and both are subject to differing courses of study. This historical dichotomy, laid out below by selective survey, has been recently challenged in cognitive science.

Action and cognition

For at least two decades, there has been a prominent movement in cognitive science to broaden the role of the body in theories of cognition (e.g., Ballard, Hayhoe, Pook, & Rao, 1997; Barsalou, 1999; Clark, 1997; Dreyfus, 1972, 1992; Glenberg & Robertson, 2000; Lakoff & Johnson, 1999; Rizzolatti, Riggio, Dascola, & Umilta, 1987; Varela, Thompson, & Rosch, 1992). The movement contrasts with the traditional focus in cognitive psychology of studying cognitive processing, such as memory, language, and categorization, as in-principle separable from the perceptual and motor systems associated with them (e.g., Carey, 1985; Chomsky, 1966, 2000; Fodor, 1975, 1983, 2000; Marr, 1982; Pinker, 1997; Putnam, 1960).

For example, explaining how we process a sentence's meaning may be couched in terms of propositional structures. When the sentence "I swung the hammer" is processed, the structures that the cognitive system manipulates are thought to be language-like elements themselves, stored, encoded, represented, etc. as information structures that have been stripped of their perceptual-motor content (e.g., Pylyshyn, 2000). In other words, hearing the word "hammer," once reaching a stage that can be called word processing, involves accessing semantic feature lists such as *is weighty* or *is long* that are encoded as semantic predicates. These predicates, "language of thought" descriptors, do not carry the perceptual (the way you might sense the weightiness of the object) or motor (the way you might hold or play with it) information about the hammer -- they are just encoded feature descriptions (e.g., Katz & Fodor, 1963).

This traditional perspective is being challenged, and some argue that these "amodal" representations do not properly characterize cognitive processing. Instead, these information structures might make use of the perceptual-motor, "modal," content that is relevant to the features (e.g., a perceptual-symbol system, Barsalou, 1999). In the case of motor information, one source of evidence for this challenge has been to investigate how action-relevant variables modulate cognitive processing. Experiments conducted by Tucker and Ellis (1998), for example, show that the recognition of an artifact (e.g., a coffee mug) is influenced by responding conditions (e.g., with the right hand) when the handle of the mug is compatible with these conditions (i.e., it is recognized faster). A similar effect is seen in sentence processing. Glenberg and Kaschak (2002) demonstrate that responding conditions (e.g., pulling towards or pushing away) can influence judging the sensibility of sentences containing compatible descriptions (e.g., *close the drawer*). These findings (along with others reviewed below) support the position that motor information is somehow being incorporated into cognitive processing -- if that were not the case, the actual characteristics of the required action would have no bearing whatsoever on the speed or reliability of the decision. Both of these examples reveal how the parameters of action can feedback into the cognitive processing relevant to the experimental task.

Such results are interesting because action is often assumed to be the point at which cognition stops. From a traditional perspective, understanding the word "hammer" works by amodal information processing, and depending on the required task response, executes a particular action the characteristics of which will not have any bearing on processing, nor contain anything resembling processing itself. For regal cognitive processes such as memory, language, and categorization, action can be seen as relatively irrelevant. The movements marshaled by motor processes are just a minor “design issue.” Hook up any machinery you want, its functioning will be slave to the decisions made by these processes.

Despite the growing evidence adduced by the perceptual-motor movement, the reigning attitude is just that: Action is not much relevant to cognition. The upshot of this attitude, only a slight caricature of the theoretical status quo, is that action phenomena have become a lesser-studied aspect of human psychology. Why would decades of theorizing in cognitive psychology proceed without regard to output processes? Recently, Rosenbaum (2005) has proposed six hypotheses that may account for the relatively low status action has in psychology. Two of these hypotheses are particularly apt here.

The first he dubs the “too-hard-to-study” hypothesis. One possible explanation for the lack of attention is that action and motor control are too complex and variable to study at any level beyond the barest detail. Rosenbaum dismisses this explanation -- the history of psychology reveals innumerable cases in which sufficient curiosity drove innovation. This innovation opens both technical and conceptual routes to understanding phenomena that interest experimentalists.

The second hypothesis, “think-before-you-act,” is one that Rosenbaum feels more successfully accounts for the neglect: All interesting psychological phenomena have come and gone by the time action is executed. He notes that perception and cognition enjoy a higher status than action because they are more often associated with intelligent behavior. Perceptual processes like visual object recognition, or cognitive ones like categorization, can presumably be studied without any thought of their associated motor accompaniments.

While he sees this hypothesis as a more plausible explanation for neglecting action, he contends that action has served many important functions in the study of both perception and cognition (such as embodied sentence processing, described previously). The short shrift is thus undeserved.

It is likely that both hypotheses explain some of the neglect of action in cognitive science. Both are historically tied to two closely related assumptions about how we decide on the subject matter of psychology, and the means by which it is studied. The basis for “think-before-you-act” is likely born of a very long tradition in western philosophy and science about the nature of the human faculty for thinking. It has provided entrenched but sometimes implicit approaches to specifying the subject matter of psychology. These approaches have placed emphasis on behavior that appears irrelevant to its precise physical manifestation, and can supposedly be studied independently of the body through which this behavior takes place. Problem solving, for example, just happens to have a lot more to do with your cranial contents than your fingernails -- moreover, the function of cranial contents can be couched in a computational vocabulary explaining this behavior without regard to the physical instantiation of the cranial contents themselves (more on this below).

The basis for “too-hard-to-study” may be argued to be, at least in the emerging sciences of the mind, an implication of “think-before-you-act.” This second hypothesis recommends that there are only certain manners in which the subject matter can be studied. The approach typically precludes data that are *inherently* related physical properties of action (particularly because such a data source is too noisy and multifarious; see below). Rosenbaum notes a few cases in which dependent measures drawn from action reveal cognitive processes, thus substantiating that action gets a short shrift. Without a detailed consideration of his examples and their implications, however, his point may be misinterpreted as trivial: Any behavioral experimentation must involve an action-based dependent measure somehow (i.e., the observables). The influence of these "too-hard-to-study" and "think-before-you-act" bases is rather more specific -- the dynamic or "internal" characteristics of a response are considered unimportant to the study of cognition. The subject matter of psychology cannot have much to do with detailed dynamic response characteristics, because these characteristics cannot reflect the processes suggested for study by the "think-before-you-act" hypothesis. Thus, at the very point at which action is initiated, thinking is over, and the separate world of action has begun.

Below I give a selective historical survey justifying this interpretation of the theoretical status quo: The view that what psychologists study -- thinking, reasoning, remembering, etc. -- takes place before action is initiated, and can proceed even without any action at all. Thus, how psychologists study thinking cannot involve action inherently. In other words: Thinking can happen without a body. Or at least, the body is trivial, being only an encasement of the thinking mind. How did we get here? The next section reveals that this attitude emerges from a long and strong intellectual tradition. Laying out the tradition might begin with the seemingly simplest question about cognitive science's subject matter: Where does thinking take place?

The "wheres" and wares of thinking

With little hesitation, most people would point to their head if asked "Where does thinking happen?" Notwithstanding the occasional humor about individual differences in this location, sincere respondents would supply this answer even when they have little familiarity with the relevant sciences. Discoveries of brain function are now comfortably tucked into the broader public knowledge about cognitive function. Such knowledge is perhaps taken for granted, given that thought was once thought to take place in the heart (French, 1978; but maybe it has something to do with it, Clark, Naritoku, Smith, Browning, & Jensen, 1999), and the brain was once thought to be a sophisticated cooling device (e.g., Aristotle's "radiator theory"; maybe it *was*, Fialkowski, 1986).

This question about the "wheres" of thought is possibly the easiest of all to answer. The "whats, hows, and whys" are considerably more complicated. This is not just because locating is a more intuitive or concrete task -- the answer also comes from a strong basis in evidence. Paralysis does not extinguish thought. When even the senses are drastically diminished, there can be normal capacity for thought, or sometimes a seemingly greater than normal capacity (e.g., Keller, 1903). We can talk ourselves through a problem without seeming to move a muscle. These and many other observations naturally motivate the conclusion that the wider body does not harbor thinking. Such a conclusion is fairly obvious when intuitive assumptions about the (perhaps) more important questions are adopted -- the whats and hows and whys of thinking. These more troublesome questions have answers that might influence the response to where thinking happens. What thinking is, how it happens, and why, seem to have a certain epistemological primacy over the wheres of thinking. Once one has some of their answers, they may recommend differing emphases on cranial contents, and their host body.

Probably the most prominent and so-far successful approach to *what* thinking is derives from the metaphor of a computing machine. Some have incorrectly asserted that this idea is unfashionable (e.g., Pinker, 1997). It is in fact born of a long and influential intellectual lineage. The notion that a logical computing machine characterizes human thinking has, some would argue, changed somewhat in form, but mostly in content, since at least Aristotle (e.g., Barendregt, 1997; Bochenski, 1951; Lukasiewicz, 1951; Rayside & Kontogiannis, 2001). Aristotle's and other ancient ideas of human reasoning look very much like the propositional structures manipulated by an information-processing module of recent conception (Fodor, 1983; Pinker, 1997; Tooby & Cosmides, 1992). Given the Hellenic roots of symbolic logic, and the central role of symbolic logic in the emergence of computing, it should come as little surprise that such a system as the Aristotelian syllogism is not far from production rules employed by symbolic cognitive models (e.g., Anderson, 1993; Schank & Abelson, 1977). These ancient systems of logic waited for thousands of years until Boole and others (Kneale, 1948) to find an explicit formalization. This symbolic logic, with its discrete structures and operations, has become perhaps the central focus of information-processing approaches to cognitive processes (Anderson, 1980; Anderson & Bower, 1973; Bechtel, Abrahamsen, & Graham, 1998; Neisser, 1967, 1976; Newell, Shaw, & Simon, 1958). The influence of Aristotle and other ancient logicians thus served to frame the nature of human reason. This tradition, ancient and robust, thus underlies the current computational approach to the mind.

"Classical" cognitive science was born of this metaphor for thinking. Both conceptual and technical innovation in the twentieth-century permitted the development of systems whose behavior resembled human reasoning

(Newell, Shaw, & Simon, 1958). For example, Newell and Simon (1956) developed a famous system that proved logical theorems, sometimes more elegantly than humans (Bechtel, et al., 1998). This system would not have been possible without the conceptual contributions of Turing and von Neumann (Anderson, 1980; Bechtel, et al., 1998), and their subsequent application in implemented computing systems. These intelligent computing systems, in a very real sense, are the progeny of this ancient conception of human reasoning, one that has existed uninterrupted from Aristotle to Chomsky. One could argue that it has always been the reigning metaphor, turned into science with the advent of computing systems, and cognitive science (see Smith, 1991, for a collection of relevant papers).

This line of thinking about thinking has often accompanied another prominent line of thought. Another related historical trend concerns the role of the body and environment underlying thinking: They are of limited importance. A prominent and again Hellenic example of this is Plato's theory of forms. The central theory of Plato's holds that the world of perceptual objects is changing and illusory -- that some other realm of objects, the realm of immutable eternal forms, is the only genuine reality. In fact, we cannot perceive these forms in the sensory sense, but only through operation of the intellect can we be aware of them. The realm of forms is the unchanging basis or "blueprint" for the ephemeral world of things (Edman, 1928; Stokes, 2002). As Plato's famous allegory of the cave illustrates, the realm of forms casts shadows into the perceptible world, which are imperfect representations of these eternal forms. The eyes are imprisoned by these worldly representations, but the intelligent mind can escape them by recollection of forms or ideas. Imagine the back of a cave, towards which numerous prisoners are forced to face. Light from a fire, behind the prisoners, is cast onto this cave wall, and between it and the prisoners puppeteers act out sequences of events that the prisoners can recognize. The prisoners mistake the shadows for the true realm. There is a realm representing the actual nature of the shadows, and the shadows are mere imperfect projections from this realm onto the cave wall.

This allegory starkly captures some modern intuitions of the role of the body and environment in cognitive processing. Consider for example Chomsky's well-known linguistic dichotomy, now referred to as E-language and I-language (Chomsky, 1986). E-language refers to language as manifested in the world of linguistic experience -- it is messy, filled with false starts, grammatical errors, and so on (Chomsky, 1965). I-language, by contrast, is our unsullied implicit knowledge of language. I-language starts its life as a universal and genetically prescribed device that can sift through the messy external language, and descend onto a stable state (which is simply a specified variation of the universal initial state, or universal grammar). The relationship between Plato and Chomsky is no coincidence. Descartes, another thinker who emphasized the same limitations of sensory experience and the role of some immutable creator and the mind's capacity for his (or her) detection, is the namesake of one of Chomsky's well-known volumes (Chomsky, 1966).

Chomsky's arguments that our language faculty is an independent and innate subsystem contributed to a powerful theoretical trend in cognitive science. The trend is best illustrated by Fodor's influential framework for cognition (Fodor, 1983), which has served as a powerful theoretical guide for the developing years of cognitive science (Applebaum, 1998). The trend in fact imports the independence of language (from the world of experience or performance) into other cognitive architectures: The whole mind actually consists of distinctly separable subsystems that not only can be studied independently, but in fact operate largely independently except for a few conveniently designed interfaces. Fodor now notes that there are inherent limitations to this perspective, but contends that it remains the most successful and hopeful perspective on cognition (Fodor, 2000), and numerous researchers have pressed its usefulness (Carey, 1985; Marcus, 2001; Pinker, 1997; Tooby & Cosmides, 1992).

There is a strong relationship between these two trends -- thinking as computation, and the realm of perceptual-motor experience as ill-suited for it. By conceiving of thinking as proceeding through logical computation, systems have long been couched in terms of language-like symbols whose origin, such as through learning, seems difficult to explain. Some have suggested radical nativist proposals about semantics to accommodate this (e.g., Fodor, 1981). The realm of perceptual-motor experience seems quite distant from these informational structures -- and are relegated to “transductor” and “effector” processes that don’t inherently participate in central processing (Fodor, 1983).

What is thinking, then? The long intellectual tradition surveyed selectively here shows a prominent role for discrete, logical computation over symbols and their relations. In addition, this nature of thinking recommends a lesser role for perceptual-motor processes -- the cognitive “realm of forms” is discretely bounded by its input and output, processes that independently sort out the messy realm of experience. The brain is now known to be the wheres of thinking. It now provides the physical substrate of the cognitive realm of forms, housed in a body whose functions are slavishly devoted to this central realm.

Recasting the “whats” and “hows” of thinking

When distinguished discretely from the bodily input-output machinery, the computational procedures underlying cognitive processes give way to the property of multiple realizability: Computational procedures are expressed in descriptive systems that can be realized in any hardware. This *functionalist* approach to cognition -- conceiving thinking *just as* the appropriate mediation of input and output (Putnam, 1967) -- does not lend itself to concern with the precise physical substrate of that input-output mediation. As long as the causal relations are preserved between internal cognitive states so as to generate reliable input-output relations, it in fact does not matter how they are physically realized. As oft-noted, thinking as computation can occur in silicon-based or carbon-based creatures, or perhaps, even instantiated by entire populations of individuals (Block, 1980, who uses this example to argue against functionalism), or perhaps exhibited at an excruciatingly long time scale by trees and plants (Dennett, 1996).

Functionalism has been frequently challenged (e.g., Block, 1980; Churchland, 1986) -- very often under the guise of debates regarding “symbol grounding.” One of the most famous critiques is Searle’s well-known thought experiment (Searle, 1984), in which a functionalist conception of intelligence is shown to be unintuitive. Another is to note that while multiple realizability of may be true of abstract computation generally, it becomes false when we settle on a more specific model system (e.g., human cognition, chimp cognition, etc.; see also Bitterman, 1960, who argues that such a specific approach is ideal in animal learning and behavior). In other words, human cognition *is* physically realized, and information-processing in humans *can be* identified with particular physical processes (Bechtel & Mundale, 1999). For example, the operation of the visual system can be understood in terms of simple and more complex feature processing in *specific* neural hardware.

Another way of accomplishing a critique of functionalism is to recast the whats and hows of thinking, placing them not in an abstract computational framework that hopes to capture intelligent human behavior as we see it fully developed in adulthood, but rather in a framework that visualizes the evolution or development of this behavior. An evolutionary perspective is perhaps particularly effective in accomplishing this recasting. Braitenberg's (1984) celebrated thought experiment helps with this.

The thought experiment works this way. Imagine designing a very simple machine or vehicle, equipped with, to begin, a sensor and a motor. The sensor is designed to activate the motor in the presence of some stimulus, such as a light. As a consequence, the vehicle will move towards this light source.

Imagine adding two sensors, one on either side of the front of the vehicle (much like headlights), and two motors (much like back tires). The left sensor is attached to the left motor, and the right sensor and motor are attached (one-to-one parallel connections). In the presence of light, the vehicle will now move away from the stimulus, since the closest sensor-motor pair to the stimulus will cause the vehicle to veer away.

Now imagine making further modifications to the vehicles. For example, try moving the left motor to the right side (and the right to the left side), while maintaining their original connections (this could be accomplished by crossing the connections, instead). This time, the closest sensor will cause the opposite motor to activate, bringing the vehicle closer to the light.

Braitenberg continues in this line, making gradual modifications to the vehicles to cause them to generate more and more complex behavior. He argues that through the law of "downhill design, and uphill analysis," we might start attributing complex internal processes to these creatures. For example, the above vehicles exemplify the rudiments of fear and attraction. Addition of novel components, such as associative neural hardware, remains intrinsically intertwined in the sensory and motor machinery of the vehicles. While the internal workings become more complex, the behaviors become even more so -- and what can appear a relatively simple modification to the inner workings of the vehicles, can suddenly appear ruminative, passionate, affectionate, etc.

The resulting interpretation of uphill analysis is not necessarily a mischaracterization. In fact, Braitenberg's likely original motivation is to reveal how our own or other creatures' complex behavior may be the result of only small and simple design modifications (such as through evolution or development). This is just what such things as "desire" and "hunger" and "affection" might be: "Let the problem of the mind dissolve in your mind." (p. 1)

To make the point relevant to current discussion, we can focus on Braitenberg's initial, simplest vehicle. This vehicle, exhibiting the simplest kind of attraction to a stimulus, is composed *just* of perceptual and motor components -- and of course a needed external stimulus to generate its behavior. The system is closed *not* by an abstract rule that describes or explains the system, but is only fully accounted for by the closed system of stimulus-sensor-motor that feed into each other continuously. The "thoughts" of this vehicle are in fact nothing but the interaction between perception and motor systems in the context of the external stimulus.

While the vehicles become more complex, one recognizes that "thinking" is always the mediation of sensor and motor components, producing complicated behavior patterns. Classical computational cognitive science recommends employing a computation-based vocabulary to explain the behavior of these vehicles. This descriptive system may work. Nevertheless, Braitenberg's experiments show that *rules* are only *descriptively* employed post hoc -- through the uphill route of analysis -- to subsequent vehicles. The *matter of fact* of subsequent vehicular cognitive processes is simply the small but important modifications to the mediation of perceptual-motor systems. The "thinking" is thus never torn from concern with perception and action. Each subsequent addition or modification must maintain how the previous instantiations accomplished this mediation.

Mental processing in the most advanced vehicles may in fact look like the ruminative behaviors of human beings. Nevertheless, because we designed these vehicles, we recognize that "thinking" is only conveniently described as a set of rules, when we know that thinking is actually the "complexification" of perceptual-motor mediation -- the internal workings may become more complex, but they are from the outset designed to appropriately generate mediation between the *specific perceptual and motor systems composing the creatures*.

This example serves to recast what "thinking" is in design terms -- or, if you like, in terms of surprisingly small modifications to an organism through the course of evolutionary or developmental change. Despite its potential intuitive power, it suffers some limitations. First, the notion of design may be a substantial simplification (Gould, 2002; but see Dennett, 1995). Secondly, it does not supply a structured argument for maintaining a deep connectivity between the body and the central processes of thought. It simply serves to "pump the intuition" (Dennett, 1988) about how the systems underlying complex behavior are ever intertwined in perceptual-motor processes. Despite this, something not unlike it has formed the basis of Brooks' well-known research program for artificial intelligence (Brooks, 1995). The thought experiment may therefore carry some strong intuitive force, while not having a complete absence of applicability.

Continuity of perceptual-motor processes and thinking: Chapters 2-5

The separation of mind and body can be saved even in this scenario by proposing predictions about the evolution of the internal mediating processes. The separation predicts that evolutionary changes led to centralized thinking processes that become gradually modularized, and independent of input-output mechanisms. Thus in the Braitenberg thought experiment, one may simply say that, through a sufficiently large number of vehicle generations, we may obtain a modular architecture, in which the original sensors and motors are discretely distinct from central processing. The human cognitive system could be such an architecture.

But this recasting of what thinking is, and how it occurs in the real world and changes over evolutionary time, suggests another radical perspective on information flow from perceptual processes into motor action. If one accepts a strong interpretation of Braitenberg's experiment, the vehicular evolution never "severs" perception and action from cognition.

Adjustment of the vehicles' perception-action mediation may develop centralized thinking processes that remain inherently tied to information about perception and action. Thus, subtle characteristics of the input to a vehicle "echo" all the way into motor processes. The idea behind this has been called the "continuity of mind" (Spivey, 2006), and is further supported by a consideration of the very machinery on which our cognitive system is based: The dynamic flow of stimulus energy (e.g., Gibson, 1979), the dynamical and fractal nature of neural firing patterns (e.g., Teich, 1989), and the complex interplay with continuously changing motor output (Kelso, 1996). Cognition might live in this continuity, and the idea that the body's input and output processes are severed from the cognitive processes is a fiction that has so far conveniently served explanation, but is of only limited importance as we focus on the continuous, temporal, dynamic nature of cognitive processing.

The next four chapters of this work seek to support this radical suggestion of the continuity between perceptual-motor processes and cognition, even in a relatively complex cognitive process. In other words, even in a Braitenberg vehicle that has reached a very high level of sophistication, the cognitive process generates motor movements that reflect the internal processing itself: Action has not been severed completely from the internal states of the cognitive process. The subsequent chapters do this by focusing on the flow of information from cognitive processing into action. Chapter 2 shows that a high-level cognitive process like categorization has this characteristic. Animal exemplars of either very high (e.g., cat) or low (e.g., whale) typicality are categorized into their superordinate classes (i.e., mammals). In previous categorization research (reviewed in Chapters 2 and 4), atypical animals result in slower, less efficient categorization, while more typical animals are quite readily categorized. This graded nature of categorization is explored in a very simple task that tracks manual output through mouse movements. Even after the action has been initiated, four experiments in Chapter 2 reveal that in both pictures and words of animal exemplars, motor output itself is graded in this manner: The typicality-based effects of categorization are also exhibited in action processes.

In Chapter 3, some discussion about the role of models in theories is discussed, and provides some theoretical justification for developing a computational model of this cognition-action flow of information. Chapter 4 bears this out, and presents two simulations which seek to articulate what kind of relationship between cognition and action can produce such graded motor output. A prediction generated by the second simulation is substantiated in Chapter 5. Typicality gradients among large classes of animal exemplars are also revealed in motor output, showing that the flow of information into action exhibits these finer-grained patterns of cognitive processing.

The implications of this research address the theoretical discussion presented above. Because cognitive processing is being reflected in action processes, it seems that some version of the continuity of mind thesis is the "matter of fact" of cognition. The separation of mind and body, while so-far useful for a variety of processes, cannot be true under deeper scrutiny of the temporal dynamics of perception, action, and cognition.

Matters of fact and units of analysis: Chapters 6-7

The final two chapters of this work aim to evaluate the nature of this debate more broadly. While computational conceptions of cognition, and powerful intuitive arguments about continuity, have both been compelling to varying numbers of researchers over the decades, the matters of fact of cognition are relevant to a very broad range of domains, from memory to language to perception. This complexity precludes a final decision about which system best captures cognitive matters of fact. The last two chapters consider how conciliation or competition between descriptive systems may be better explored.

Chapter 6 lays out the area of symbolic dynamics for conceptual purposes in cognitive science. While categorization may exhibit this gradedness, the discrete and serial theoretical framework continues most successfully to address other higher-order cognition processes (e.g., problem solving). This chapter showcases the mathematical framework of symbolic dynamics as a potential means of reconciling debate between these theoretical extremes. This set of tools, used in a variety of physical and mathematical sciences, offers a formal terrain that can incorporate both discrete and continuous accounts of cognitive processing. This mathematical framework may serve to formally adjudicate between them, or indeed reveal that both kinds of representation are important for explaining cognitive processes.

Chapter 7 concludes the dissertation with an extensive discussion of the philosophical and theoretical implications of both the research presented in Chapters 2-5, and the proposals for symbolic dynamics in Chapter 6. I consider a few ways in which the "matters of fact" may reach a consensus.

Three such possibilities are discussed, guided partly by issues in the philosophy of science. First, the matters of fact of cognition will always be “underdetermined” given any or all the evidence available. In other words, there will always be disputes about how cognition actually, in reality, works. We will not enjoy the satisfaction of a consensus. The second possible outcome is that cognitive science will be happy with the plurality of models -- both in number and kind -- in its various subdomains. This contentment will permit matters of fact to simply be relevant to the successful model in a given domain. Such a “metaphysical pluralism” urges that matters of fact are obtained across a whole range of domains of inquiry. The third draws on Chapters 2-5, and argues for continuity based on an “enhanced reductionism.”

Instead of embracing any one of these approaches, I will argue that all three are not very satisfying. I present a fourth option in which this debate regarding matters of fact can be reduced to units of analysis, and the goals of a particular analysis over those units. Again drawing on discussion in the philosophy of science, I will be somewhat dismissive of matters of fact, and propose a pragmatic solution. The goals of cognitive science are not to discover matters of fact, but to solve particular problems within particular domains. The upshot: Matters of fact don’t matter.

CHAPTER TWO

Graded Motor Responses in the Time Course of Categorizing Atypical Exemplars

Introduction

The past few decades have seen a transition from classical set theoretic accounts of categorization, wherein cognitive processes and categories are seen as discretely bounded with unique membership, into the development of nuanced theories of fuzzy categories and their interrelationships. For example, Reed (1972), Rosch (1975; 1973), and Rips, Shoben, and Smith (1973) challenged the classical conception of category structure early on, revealing that more graded semantic-space effects are readily observable, and not easily accountable in terms of classical conceptions (e.g., Bruner, Goodnow, & Austin, 1956; Collins & Quillian, 1969). The emerging probabilistic prototype framework has also faced challenges from alternative accounts. An exemplar-based theory of categories is often argued to account for a wider set of data (Medin & Schaffer, 1978; Nosofsky, 1988, 1992; Medin & Ross, 1989; see recently, Storms, 2005; see Smith, 2002, for an opposing perspective). Aspects of these similarity-based accounts, both prototype and exemplar, have also been suggested to figure into a theory-based theory of category structure in our cognitive system (Medin, 1989; Murphy, 2002; Murphy & Medin, 1985; see, e.g., Heit, 1994; Burnett, Medin, Ross, & Blok, 2005 and Lynch, Coley, & Medin, 2000). Despite this multiple branching of theoretical directions, experimental methodologies had remained largely unchanged. The upshot, until recently, is that the *time course of processing* in categorization had been underexplored, compared to the development of theories regarding the static representation of category knowledge.

There is a growing body of research devoting itself to this question. The time course of categorization phenomena has begun to receive some attention, from perceptual categorization (Ashby, Boynton, & Lee, 1994; Lamberts, 1995, 2000; Nosofsky & Palmeri, 1997), to categorical perception of speech (McMurray & Spivey, 1999; McMurray, Tanenhaus, Aslin, & Spivey, 2003) and lexical processing (e.g., Cree, McRae, & McNorgan, 1999). Most of these experiments have made use of reaction-time measures. For example, Lamberts (2000) introduces an information-accumulation account of speeded classification of objects (see also Ashby et al., 1994; Nosofsky & Palmeri, 1997). These investigations of the time course of categorization seek to supplement research that has generally placed more emphasis on the outcome of the process, rather than its continuous temporal dynamics (Nosofsky & Palmeri, 1997). Decision models of categorization, like Lamberts' (2000, 2002) and Nosofsky and Palmeri's (1997), uncover the time course of the process leading up to the button-press response.

It is possible, however, that the motor output of the decision process itself also exhibits graded effects over time, providing further information about the continuous nature of the process. For example, McMurray et al. (2003) used eye-movement data to investigate the graded temporal dynamics of speech sound classification over the course of several hundred milliseconds. Similarly, Nederhouser and Spivey (2004) used eye-movement patterns to show competition between alternative taxonomic classifications of atypical animals. When given a toy whale to categorize by dropping it in either the "fish" bucket or the "mammal" bucket, participants frequently fixated the “fish” bucket first, before then fixating the “mammal” bucket for guiding the hand movement. These kinds of eye movement data represent “micro-decisions” that are tentative intermediate emissions from a categorization process that is extended in time (Richardson, Dale, & Spivey, in press; see also Rehder & Hoffman, in press).

Motor responses of this kind epitomize what is often considered “output” from cognitive processes, the outcome of a pipeline from perceptual processes, through association cortex, into premotor regions, then collapsed into individual manual and oculomotor choice behaviors. More recent work on manual and oculomotor movements suggests that these manual processes are contiguous with cognitive processes (Gold & Shadlen, 2001; Shin & Rosenbaum, 2002). The eye-movement patterns themselves offer an unusually early glimpse into partially-active “micro-decisions” in motor output (cf. Gold & Shadlen, 2000; Magnuson, 2005). Unfortunately, the ballistic quality of most saccades prevents them from being able to exhibit truly graded effects of partially activated representations. On any given trial, participants either fixate the competing object or they don’t. The eye movement data are usually not able to show continuous attraction effects within a trial (but cf. Doyle & Walker, 2001). In contrast, non-ballistic arm movements regularly involve curvature that reveals continuous attraction effects (Goodale, Pélisson, & Prablanc, 1986). In fact, continuous manual motor output has been studied for the very purpose of providing a variety of clues about graded underlying cognitive processes within individual trials. For example, Abrams and Balota (1991; see also Coles, Gratton, Bashore, Eriksen, & Donchin, 1985) used an arbitrary manual response task (pulling a handle to the right or left) in a lexical decision and recognition memory test. The frequency of lexical items and strength of recognition memory were strongly associated with a shorter time spent pulling the handle to its limit (i.e., greater force and velocity of pull), independent of the latency of response onset. The authors recommend that anticipation components of response preparation and force and velocity parameters of motor execution must figure into our understanding of cognitive processing. Similarly, Spivey, Grosjean, and Knoblich (2005) used computer-mouse trajectories to study spoken word recognition. In individual trials, subjects heard instructions such as “Click the candle,” and selected one of two objects in two corners of a computer monitor. Trials in which the two objects had similar-sounding names (e.g., a candle and a candy) revealed mouse-movement trajectories that exhibited significant attraction toward the competing object, compared to control trials (e.g., a candle and a spoon). Computer-mouse trajectories thereby revealed continuous dynamic partial activation of multiple competing representations during real-time spoken word recognition (see also Allopenna, Magnuson, & Tanenhaus, 1998, for related eye-movement results).

The time course of categorization is a suitable model system for understanding the interaction among different processes as cognition unfolds (e.g., memory and attention; see Lamberts, 2000, for a discussion). For example, feature-sampling models of the time course of categorization contribute to understanding not simply categorization outcomes, but also predictive or explanatory systems accounting for reaction-time measures in these studies, and how these measures reflect perceptual information accumulation used in categorization (e.g., Lamberts, 2000, 2002). In the present case, four experiments reveal that categorization as a cognitive process flows into, rather than collapses onto, the effectors responsible for manual action.

The goal of the current paper is to demonstrate, via continuous computer-mouse trajectories, that graded representations of this kind are also involved in the real-time taxonomic categorization of animal names and animal pictures. It is suggested that the temporal dynamics of matching a category exemplar to one of its potential categories exhibits a degree of "granularity" (Miller, 1982; Miller & Ulrich, 2003) that may be problematic for a discrete representational account of the categorization process. We present four experiments tracking continuous manual motor output during lexical and perceptual categorization of atypical exemplars. This work adds a new methodology to the study of the time course of categorization, and contributes to further specifying the processes by which the mind settles onto one categorical response versus another. These experiments illustrate how the gradual accumulation of evidence for a given category is not solely composed of that category's representation transitioning from zero activation to full activation. As in many complex dynamical systems, there is also competition from alternative category representations that are partially active at the same time. Much like a dynamical system continuously traversing its high-dimensional semantic space over time, the trajectory of the categorization process reveals that multiple categories (attractor basins) are nearly visited as the system eventually settles into a unique outcome-based response.

Experiment 1

Typicality has become one of the most thoroughly studied aspects of categories and concepts since the transition from classical perspectives (Medin, 1989; Murphy, 2002). The empirical drive towards non-classical accounts of categories was guided by studies of membership and typicality judgments and corresponding reaction-time measures (e.g., McCloskey & Glucksberg, 1978; Rips et al., 1973; Rosch, 1975). Category members deemed more typical are recognized faster (Rips et al., 1973), more consistently (McCloskey & Glucksberg, 1978), have many features in common (Rosch & Mervis, 1975), and can even result in facilitated language comprehension (Garrod & Sanford, 1977; see Murphy, 2002, chapter 2, for a review of these and other robust results). Recent research on categories and concepts has continued this emphasis (e.g., Burnett et al., 2005; Estes, 2003; Murphy & Ross, 2005; Op de Beeck and Wagemans, 2002; Smith, 2002; Verbeemen, Storms, & Verguts, 2003). For example, Burnett et al. (2005) revealed that typicality ratings of various fish species are guided by knowledge-level expectations regarding fish, rather than a simple notion of category centrality (see also Lynch, et al., 2000). Smith (2002) recently used expected patterns of typicality across category members as a focal point for theoretical debate, such as the effectiveness of exemplar-based perspectives. Typicality is thus central to our categorizing capacities, but also useful as a crucial measure for comparing opposing theories.

In the following 4 experiments, we utilize an animal stimulus set (words and pictures) whose elements differ in typicality. We used a small stimulus set of well-known animals that are either highly typical members of a category (*dog* as *mammal*), or those widely regarded as atypical (*penguin* as *bird*). In this first experiment, we recorded continuous manual responses in a categorization task. Participants performed a simple judgment: After seeing an animal word, participants clicked on one of two category labels to which the animal belongs. Some trials involved animals of an atypical nature – such as a whale, which has several properties that suggest another category label (*fish*), potentially causing competition with the correct label (*mammal*). Just as in Spivey et al. (2005), echoes of a competitive categorization process should be reflected in continuous mouse movements. In other words, when categorizing *whale* as *mammal*, participants' mouse trajectories should gravitate toward the competing category (*fish*) more so than when categorizing a typical exemplar label, such as *cat*.

Participants

41 undergraduate participants signed up for this experiment for extra credit in their psychology classes. All participants in this and subsequent experiments used their right hand to perform the task.

Materials

Word stimuli were presented using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were faced with the task of choosing the appropriate category for each animal word, presented in text on the monitor. These categories included pairings of the following: mammal, reptile, bird, fish, amphibian, and insect.

There were two within-subject conditions. In control trials, the animal words were typical category members (e.g., *cat* as *mammal*). In the experimental trials, they were atypical category members (e.g., *penguin* as *bird*), and both the correct category name and the featurally-similar competing category name were presented (*bird* and *fish*, respectively). Although we assume that typicality of category membership is a graded parameter, the two-condition design of these studies required that we keep within-group variation among stimulus items to a minimum (particularly in the atypical animal condition). Therefore, only 6 highly atypical animals were used in the atypical condition, with other merely moderately atypical animals being excluded from the design. As typical category members are easier to generate, 13 highly-typical animals were used for the control condition (see Table 2.1).

Procedure

Participants were presented with two different animal category names, randomly assigned to one of the upper corners of a computer screen. After 2000 ms, a moment to see the category options, the text “Click Here” appeared in the bottom center of the screen. The participants were instructed to click first on that text and wait for an animal word to appear in its place, then to click on the upper (left or right) category name that was appropriate for that animal. Participants were provided with three practice trials before beginning the 19 target trials. All trials were presented in random order.

It was predicted that, in the experimental trials, mouse movement trajectories would show evidence of competition between the categories. This competition is revealed by analysis of mouse-movement trajectory divergence: Atypical animal trials should have movement trajectories that reveal a slight bias towards the competing category (e.g., with *whale*, a slight attraction toward the category *fish*) when compared to control trials. Data for testing this prediction were collected by recording x and y coordinates of mouse-movement trajectories. Due to occasional skipped samples, PsyScope’s sampling rate averages approximately 42 Hz. As a result, each trial collects about 40-80 mouse-position data points.

Data Analysis

Numerous analyses were conducted on these rich trajectory data. First, to enable averaging of full trajectories from multiple trials, all trajectories were normalized to 101 time steps, and were translated to begin at an x,y coordinate of (0, 0). These "time-normalized" trajectories can be compared between typical and atypical conditions. Divergence of the two averaged trajectories is established by significant differences between the x-coordinates.

This first analysis provides information regarding the overall shape of the trajectories in both trial types. In addition to this time-normalized analysis, a "space-normalized" analysis was run in which beginning and end coordinates of each trial were normalized to (0, 0) and (1, 1), respectively. Real time information was retained by computing x,y coordinates as they travel from 0 to 1 in time bins of 0-500, 500-1000, and 1000-1500 ms. This provides a window onto the movement in real time from start to finish of a trial, and includes enough data in each time bin to permit an additional statistical test of the difference between trial conditions by subjecting these bins to a repeated-measures ANOVA.

Finally, a number of properties of the trajectories were computed and compared between conditions: mouse-movement initiation time, movement duration, total categorization response time, distance traveled in pixels, and direction in degrees of the first pair of mouse movement samples. This last measure, initial movement direction in degrees, offers insight into early stages

Table 2.1: Atypical and typical animals/words in Experiments. In parentheses are the response options given to the participants, with correct categories italicized. Bold options in atypical trials indicate non-competing labels used in Experiments 2 and 4.

Atypical

Eel (fish; reptile; **mammal**)
Sea lion (mammal; fish; **reptile**)
Butterfly (insect; bird; **reptile**)
Whale (mammal; fish; **bird**)
Penguin (bird; fish; **mammal**)
Bat (mammal; bird; **reptile**)

Typical

Cat (mammal; reptile)
Horse (mammal; bird)
Salmon (fish; mammal)
Rabbit (mammal; reptile)
Hawk (bird; reptile)
Sparrow (bird; mammal)
Chameleon (reptile; insect)
Dog (mammal; insect)
Goldfish (fish; amphibian)
Alligator (reptile; mammal),
Lion (mammal; fish)
Rattlesnake (reptile; amphibian)
Shark (fish; mammal)

of the trial. For example, it is possible that participants are strategically moving the mouse cursor along the vertical on all trials before turning in the direction of the target category. Angle information on initial movement will reveal whether this is the case. Such a strategy would cause that initial movement from (0, 0) not to be significantly different from 0 off the vertical in both typical and atypical conditions.

These final measures were supplemented by two further analyses: a) tests of bimodality in the distribution of trajectory curvatures, and b) a time series analysis to explore the complexity of the resultant trajectories. As discussed in Spivey et al. (2005), any pattern of competition seen in atypical trials may in fact be the result of an averaged bimodal distribution. If half of the correct trials involved movement straight toward the target, and the other half involved movement straight toward the competitor, followed by a corrective movement toward the target, the average of all trials could appear as a graded curvature toward the competitor – when in fact it would be better described as the result of a number of trials with rather discrete errors that were corrected mid-flight. To approach this problem, each trial's area (in pixels) was computed between the actual trajectory and a straight line connecting the start and end points. A distributional analysis of this area should show bimodality if the averaged trajectory's apparent graded curvature is actually due to occasional discrete errors that get sharply corrected partway through the movement. Finally, sample entropy (Richman & Moorman, 2000) was used to discern whether atypical trajectories are more "complex" than typical trajectories (see Feldman & Crutchfield, 1998, for a discussion of complexity measures and their advantages and problems). If the competing category in the atypical trials is acting as a substantial second attractor on the dynamics of manual output, then this stress should be evident as less smoothness in the trajectories. Atypical trials should show more complexity, because two attractor forces are acting on them. Sample entropy gives a larger value the more "complex" or irregular a time series. Further detail on this measure is presented below.

Results

Participants categorized atypical exemplars with 90% accuracy, and typical exemplars with 95% accuracy, a significant difference ($p < .05$). Only correct trials were used in the following analyses. All incorrect trials were removed from analysis in this and subsequent experiments.

Time-normalized analysis

In this analysis, *t*-tests were conducted to compare the difference between the x-coordinate values for typical and atypical trials to zero, at each of 101 interpolated time-steps throughout the trajectories (see Fig. 2.1A). Rightward and leftward responses were pooled to maximize statistical power. The relevant null hypothesis was that the difference between the atypical and typical trajectories' x-coordinate at any given time step (out of 101) should be 0. As a more conservative test, in this and subsequent experiments, a reliable divergence was defined as a minimum of 8 consecutive time slices in sequence, in which the normalized trajectory differences were significant at a criterion of $p < .05$. This criterion was established by performing a Bootstrap of 10,000 simulated experiments of the same mean and standard deviation (see Appendix). Significant divergence between trajectories is thus observed when there is a substantial sequence of consecutive significant *t*-tests between atypical and typical x-coordinates. In this experiment, the trajectories exhibited significant differences in x-coordinates for 38 consecutive time slices, from the 47th to 85th time steps (*p*'s < .05).

As an additional statistical test, we computed pooled bins from these time-normalized trajectories and conducted a 2 (typical vs. atypical) x 3 (1-33, 34-67, 68-101 step bins) repeated-measures ANOVA. This test reveals a strong effect of trial type (*F*(1,40) = 21.8, *MSe* = 1824.8, *p* < .001), main effect of bin (*F* (2,39) = 861.1, *MSe* = 1916.5, *p* < .001) and a significant interaction (*F*(2,39) = 6.2, *MSe* = 783.7, *p* < .01). To reveal what portions of the trajectory are exhibiting this divergence, we conducted planned comparisons between trial types within each bin. These show a significant difference between trial types in the second and third bins (*p*'s < .001). Thus, by the second and final third of the time-normalized trajectories, atypical categorization is exhibiting significant divergence in x-coordinate from typical categorization.

Space-normalized analysis

Fig. 2.1B shows a graph of the trajectories, from leftward and rightward movements, in separate conditions in terms of time bins and normalized pixel coordinates. While the previous analysis preserved raw space and normalized time into 101 bins, this analysis preserved real time and normalized spatial coordinates of the mouse from (0, 0) to (1, 1), and pooled these values into three time bins: 0-500, 500-1000, and 1000-1500 ms. Once again, absolute left and right values of the x-coordinate were pooled for statistical comparison. A similar 2 (trial type) x 3 (time bin) repeated-measures ANOVA was used. This revealed a strong effect of trial type (*F*(1,37) = 52.3, *MSe* = .031, *p* < .001), time

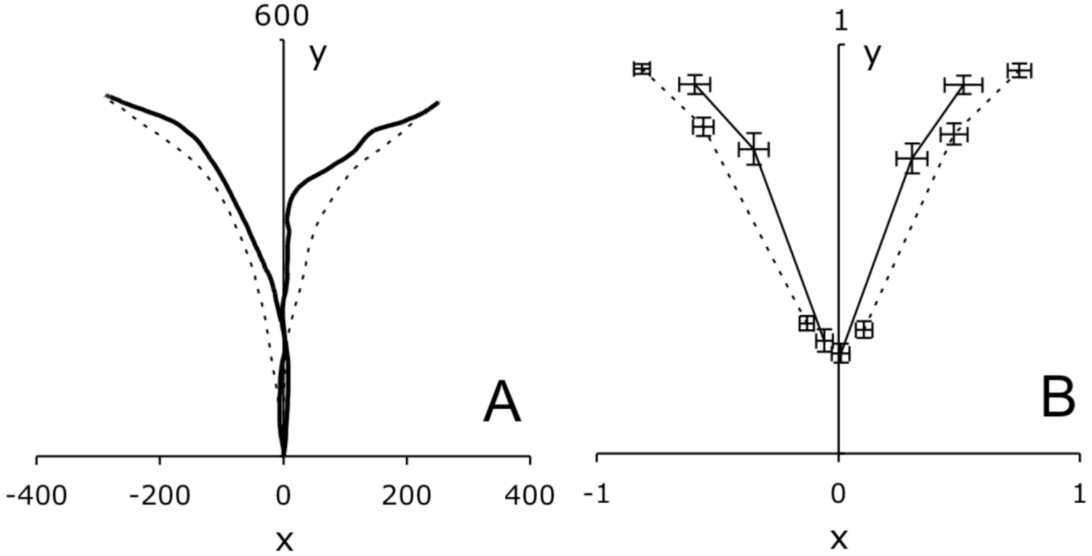

Figure 2.1: A) Mean time-normalized mouse-movement trajectories in Experiment 1, separated for left- and right-ward responses to the correct target. Atypical trials (solid lines) indicate an attraction towards the competing category on the opposite side of the computer screen in the 101 time steps. B) Mean mouse-movement trajectories in space-normalized analyses, with x,y coordinates traveling from (0, 0) to (-1/1, 1) in time bins: 0-500, 500-1000, and 1000-1500 ms. Atypical trials (solid lines) reveal slower movement towards target.

bin ($F(2,36) = 168.0$, $MSe = .041$, $p < .001$), and a significant interaction between time bin and trial type ($F(2,36) = 5.2$, $MSe = .030$, $p < .05$). Once again, planned comparisons were run between types at each time bin to detect what portion of the trajectories had significant divergence. The difference between the trial types was significant at every time bin (p's $< .01$).

Additional measures

The movement duration for atypical trials was 1573 ms, compared to 1338 ms for typical trials ($t(40) = 3.8$, $p < .001$). Total categorization time (from animal word onset to final mouse-click) for atypical animals was greater than for typical trials, with 1997 ms and 1807 ms respectively ($t(40) = 3.0$, $p < .01$). Total distance traveled also differed significantly (atypical: 894 pixels vs. typical: 831 pixels, $t(40) = 3.0$, $p < .01$). Curiously, movement initiation latency showed an opposite pattern, but was only marginally significant (atypical: 424 ms vs. typical: 470 ms, $t(40) = -1.8$, $p = .09$). An analysis of the initial direction of the mouse trajectory, from position (0, 0), extracted a measure in degrees from the vertical (from the start-click event). The typical trials exhibited a significant positive angle (toward the target) in a one-sample *t*-test (null hypothesis that degrees are different from 0), with a value of 6.0°, $t(40) = 2.6$, $p < .05$. Atypical trials in fact exhibited a slight negative angle (toward the competitor), with -.99°, but this was not significant.

Item-based repeated-measures ANOVAs were also run using these same measures. Both movement duration and distance were marginally significant in the expected direction ($F(1,17) = 3.2, 2.0$, $MSe = 72125.3, 7685.4$, $p = .09, .08$, respectively), with means of atypical trajectories slower in time and longer in distance to reach their target. Neither movement initiation latency nor total response time was significant. These results may be difficult to interpret due to the size of the stimulus set. However, the marginal significance is suggestive that the effects in the expected direction may be significant with a larger set, while the movement initiation latency is not likely to be ($F(1,17) = .8$, $p = .38$). Initial angle of movement from vertical in fact shows a marginal difference between conditions in an item-based analysis (*F*

(1,17) = 3.4, *MSe* = 54.3, $p = .08$), with the typical trials again showing the only significant difference from 0° at 6.6°, $t(12) = 3.1$, $p < .01$.

Two additional analyses were used to reveal more characteristics of the trajectories. First, as in Spivey et al. (2005), we explored the nature of the distribution of curvature across all trajectories. The pattern of divergence (from a straight line) in atypical trials could, in principle, have simply been a reflection of an averaged bimodal distribution. In contrast, there is no theoretical reason to expect such bimodality in the typical trials. The trial data used was the area between the actual trajectory and a straight line from (0, 0) to the final click (see Spivey et al., 2005). All trials across participants were used in order to have sufficient numbers for the statistical test. All area values were converted into *z*-scores, and subjected to distributional analyses. A Kolmogorov-Smirnov test for a difference in distribution revealed no difference between typical and atypical trajectories ($\chi^2 = 1.9$, $p = .79$; see Fig. 2.2). In addition, computed bimodality coefficients for both typical and atypical trials were in the unimodal range of <.555 (see Spivey et al., 2005). While typical (.536) and atypical (.535) were close to this cutoff at which concerns about bimodality may arise, the values are in fact almost identical to each other. These results indicate that even if bimodality were a concern, it is not exclusive to the atypical trials.

Finally, we explored the complexity of the trajectories using a time series analysis giving an entropy-based value. Sample entropy (Richman & Moorman, 2000) is computed by comparing windows of size *m* from a time series. A given window is said to be similar to another if their distance is less than some value *r*, known as the tolerance. This is given by the following

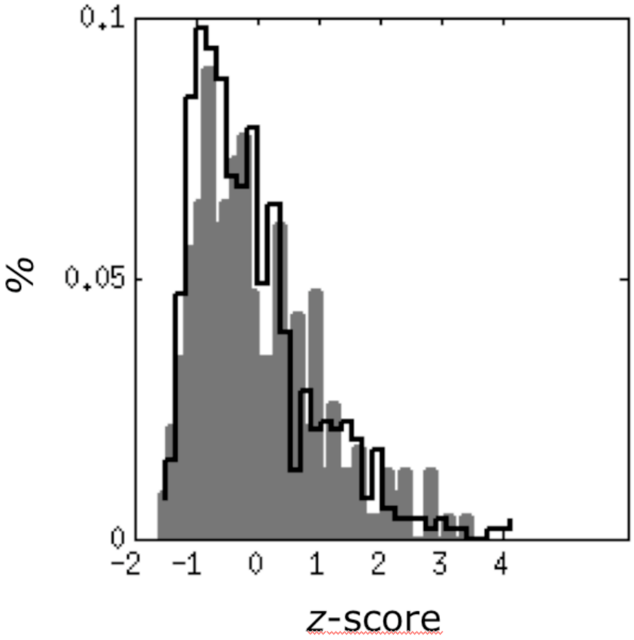

Figure 2.2: Percentage distribution histogram of *z*-scores for area in pixels in atypical (solid grey distribution) and typical (black outlined distribution) superimposed. Distribution analysis reveals no significant difference between these distributions.

equation, in which two windows (w_1 and w_2) of size *m* have a distance equal to the maximal difference between paired elements composing them, with $w_1(i)$ representing the *i*-th element of window w_1:

Sample entropy is then computed by taking the difference between averaged natural logarithms of counted similarities at size *m* and *m+1.* If, when window size is increased, the average number of similarities is similar, this value will be low. Conversely, with a higher difference between counts in *m* and *m+1*, it is less likely that windows similar to each other with size *m* will again be similar at *m+1*, indicating more irregularity in the time series. For this analysis, we used the time series of a trial's normalized x-coordinate fluctuations. This time series represents the extent to which horizontal movement is fluctuating towards one category label or back to the other. A range of window sizes was used (2-10), along with a tolerance of .2 multiplied by the standard deviation of all x-coordinate fluctuations (x_{t+1} - x_t) in normalized trajectories. Across all values for *m*, mean sample entropy is numerically greater (indicating higher complexity) in atypical trajectories than in typical trajectories. These differences are statistically significant or marginally significant when *m* has the values 3 through 6.

Discussion

Results indicate that the process of categorization exhibited nonlinear time-course effects in mouse-movement trajectories. Trajectories revealed significant attraction toward the competing category name in the atypical-animal condition, when compared to the typical-animal condition. Interestingly, the movement-initiation latency did not generate significant differences, though measures based on the manual motion itself revealed spatial divergence between atypical and typical trajectories and a difference in categorization time. We therefore conclude that the manual output from the categorization process is reflective of a cognitive temporal dynamics wherein the mapping of exemplar to category evolves nonlinearly over time (cf. Lamberts, 2000).

Experiment 2

The same animal names were used in this second experiment. However, in the atypical-animal condition, the alternative (incorrect) response option was *not* the taxonomic class that shares similar properties with the label's referent. To explore whether the graded attraction effect in computer-mouse trajectories is in part due to competition between strongly activated categories, or if the manual curvature is solely a result of uncertainty in the match between atypical animal words and their correct categorization, we altered the response options for those trials. For example, whereas the previous experimental trials used the categories *fish* and *mammal* for the atypical animal *whale*, the present experiment used an alternative (incorrect) category that did not share features with the animal word (e.g., *bird* and *mammal* for *whale*). If the mouse trajectories for the atypical trials show equal curvature toward the incorrect response option (as in Experiment 1), this would suggest that the curvatures we are observing are merely due to a slow accumulation of evidence for the atypical animal being categorized as its correct taxonomic class; if this curvature is reduced, it would suggest that the differences in Experiment 1 are due in part to a dynamic attraction effect exerted by the featurally-similar competing category.

Participants

41 undergraduate participants in this experiment received extra credit for psychology classes.

Materials and Procedure

This experiment used the same technique and same words as the first, but the atypical-animal condition no longer involved competing categories. Instead, the correct category was paired with a non-competing one (e.g., whale = *mammal* or *bird*). The control trials were the same as in Experiment 1 (see Table 2.1).

Results

Participants categorized atypical exemplars with 92% accuracy, and typical exemplars with 98% accuracy ($p < .01$). Again, only correct trials were used in the analysis.

Time-normalized analysis

Normalized trajectories revealed a single sequence of 8 *t*-tests (*p*'s < .05), from the 72nd to 79th time step, showing atypical-typical x-coordinate differences that were reliably greater than zero (see Fig. 2.3A). Although significant, this brief divergence between atypical and typical trajectories is a substantially diminished sequence compared to the previous experiment's 38 consecutive time steps with significant differences. Again, as an additional statistical test, pooled bins from these time-normalized trajectories were used in a 2 (typical vs. atypical) x 3 (1-33, 34-67, 68-101 time steps) repeated-measures ANOVA. There was no significant effect of trial type ($F(1,40) = 2.2$, $MSe = 1341.2$, $p = .14$), but a significant main effect of bin ($F(2,39) = 1147.7$, $MSe = 1289.0$, $p < .001$), and a significant interaction ($F(2,39) = 4.0$, $MSe = 617.4$, $p < .05$). While

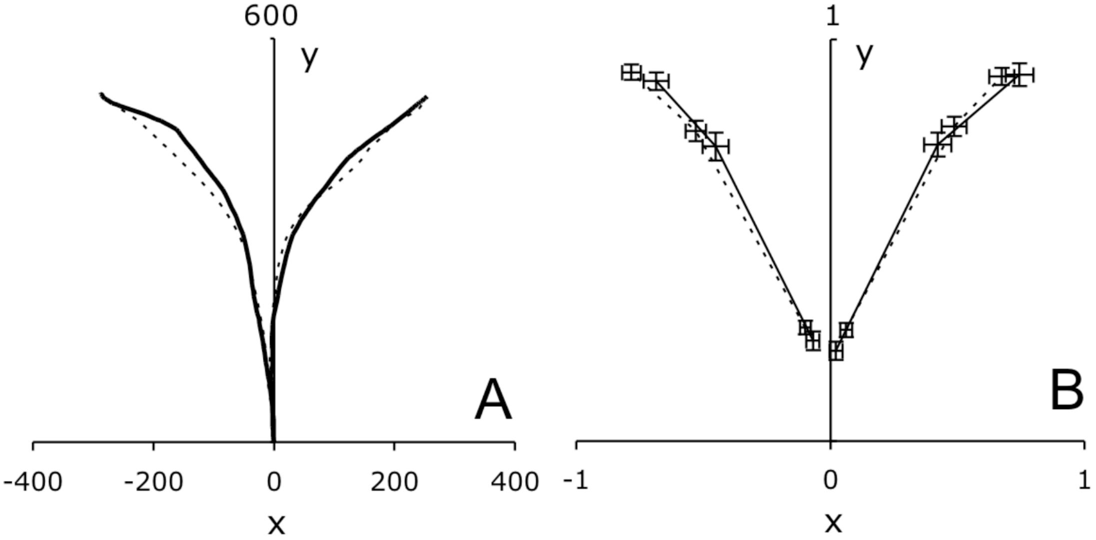

Figure 2.3: A) Mean time-normalized mouse-movement trajectories in Experiment 2, separated for left- and right-ward responses to the correct target. Atypical trials (solid lines) reveal a diminished attraction towards the competing category. B) Space-normalized time bins show no significant differences between trial types (atypical in solid lines).

there was no main effect of trial type, comparisons did show a significant difference between typical and atypical conditions at the third time bin ($p < .05$).

Space-normalized analysis

There was a main effect of time bin ($F(2,35) = 340.1$, $MSe = .035$, $p < .001$), but no significant effect of trial type, or interaction between trial type and time bin (see Fig. 2.3B).

Additional measures

No significant differences were found in movement initiation latency (416 ms atypical vs. 403 ms typical, $t(40) = .6$, $p = .6$). However, there were significant effects of total response time (1840 ms vs. 1718 ms, $t(40)=2.5$, $p < .05$), movement duration (1424 ms vs. 1315 ms, $t(40)=3.8$, $p < .05$), and distance traveled (836 pixels vs. 780 pixels, $t(40)=2.3$, $p < .05$). Initial movement angle showed no significant effects. Interestingly, both were positive towards the target category (3.1 vs. 2.25 in atypical and typical trials, respectively). In item-based repeated-measures ANOVAs, only total response time and distance traveled were significant ($F(1,17) = 4.7, 9.9$, $MSe = 20023.7, 1990.5$, respectively, p's $< .05$). Neither movement initiation latency nor movement duration was significant. There were no significant results in initial angle.

As before, a Kolmogorov-Smirnov test did not show a significant difference in the distribution of trajectory curvatures between typical and atypical trials ($\chi^2 = 1.8$, $p = .80$). Bimodality coefficients were again similar, and below .555 (.546 and .549). Finally, sample entropy analyses did not show any significant or marginally significant differences using window sizes of 2-10.

The greatly reduced trajectory attraction effects during categorization with non-competitive alternative categories suggests that the spatial attraction effects in Experiment 1 were not due merely to the inherent atypicality of the label's referent. The featural similarity between the animal and the (incorrect) alternative category – present in Experiment 1 and absent in Experiment 2 – played a substantial role in eliciting curvature in the computer-mouse trajectory. When differential divergence between the average time-normalized trajectory for atypical animals and the average trajectory for typical animals is plotted over time for Experiment 1 and for Experiment 2 (Fig. 2.4), the

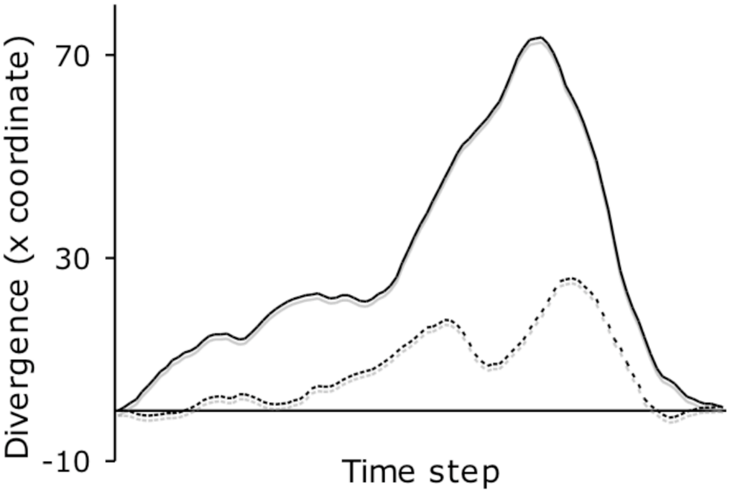

Figure 2.4: Mean total divergence (atypical trial – typical trial x-coordinates) for Experiments 1 and 2. Experiment 1 (solid line) exhibits significantly more divergence.

competitive alternative categories in Experiment 1 clearly show a much stronger attraction effect than the non-competitive alternative categories in Experiment 2. These trajectories differ significantly for 19 time slices, from the 57th to 75th time steps (*p*'s < .05). In addition, when atypical-animal trials from these experiments are compared using space-normalized time bins, there is a main effect of experiment (*F*(1,75) = 5.8, *MSe* = .096, *p* < .05) and an interaction between experiment and bin (*F*(2,74) = 3.4, *MSe* = .039, *p* < .05). These differences suggest that Experiment 1's effects were strengthened by the presence of competitor labels.

Discussion

Interestingly, and perhaps surprisingly, non-competing alternative category labels in Experiment 2 still induced some significant, albeit small, attraction. For example, when faced with possible responses *mammal* and *bird* for the examplar *whale*, participants' trajectories in Experiment 2 showed some significant attraction toward the putatively non-competing label *bird*. Why should this be? The answer may lie in the between-category similarity.

In addition to considering the match between an exemplar and a potential category, we must also consider the similarity between categories themselves (cf. Storms, 2005). Although one may initially assume that the non-competing alternative categories of Experiment 2 should remove the patterns of competition seen in Experiment 1, there actually remains substantial between-category similarity among the correct and alternative category responses. For example, in addition to *whale* being a good match for *mammal*, and a partial match for *fish*, it also has a moderate amount of fit to *bird*. Whales and birds both move, breathe, and eat. They both have eyes, skin, muscles, bones, lungs, a brain, a heart, etc. Thus, even our "non-competing" category labels in Experiment 2 still share a number of similar features with the target stimulus, since they are all in the superordinate-level category of *animals*. Even such minor similarity appears to be sufficient to produce some mild attraction effects in the mouse-movement trajectories. Based on the combined results of these first two experiments, we suggest that mouse movement curvature reflects a competition process wherein partially active categories pull the state of the cognitive system toward their respective attractor basins. Dynamic properties of the resulting nonlinear mental trajectory are emitted in the continuous manual output of computer-mouse movement. As in many classic categorization studies (e.g., Rosch, 1973; Rips et al., 1973), we were able to induce these effects with atypical category members presented in lexical form.

Indeed, the initial process of visual word recognition may also exhibit some attractor dynamics of its own (e.g., Hinton & Shallice, 1991; McRae, de Sa, & Seidenberg, 1997; Rueckl, 2002). Because initial processing in the trials of both Experiments 1 and 2 involve lexical recognition, the competitive process following them is attributable to the process of categorization. Some additional dynamics therefore take place following this word recognition phase, in the time course of mapping lexical processing onto a categorization response. These experiments reveal that even the manual output of this mapping reflects the partially active representations underlying it.

Experiment 3

In the next two experiments, participants categorized *pictures* of our animal set, rather than lexical items. Numerous studies have considered the distinction between processing pictures and words, and the relationship between this processing and a proposed core conceptual representation underlying them. Early on, for example, Snodgrass (1984) argued that pictorial and verbal codes have separate functions and distinct neurophysiological realization, rather than being mere surface forms for a central underlying concept representation (see also Paivio, 1991). More recently, numerous studies have explored the efficiency with which picture vs. word stimuli are categorized, with most showing a small but significant advantage for pictorial stimuli (e.g., Job, Rumiati, & Lotto, 1992; Snodgrass & McCullough, 1986; Viswanathan & Childers, 2003). For example, Viswanathan and Childers (2003) demonstrated a processing advantage for pictorial stimuli in a task where participants judged whether two stimuli were members of the same category. The authors suggest that visual stimuli induce simultaneous processing of category-relevant visual features *and* of their categories or concepts. Thus, the accumulation of featural-match information may be faster and more robust with pictures than with lexical presentation alone, which relies only on conceptual information acquired through orthographic input.

This discussion suggests a few predictions for Experiments 3 and 4, using visual stimuli. First, typical-animal picture trials will likely be categorized faster and more efficiently than typical-animal lexical trials. Previous research suggests this processing advantage for visual examplars. In addition, because visual stimulus features simultaneously act with conceptual activation, there may in fact be more competition for atypical picture trials (Experiment 3) than there was for atypical word trials (Experiment 1). Visual features of atypical animals often provide some partial match to competing categories. For example, whale images offer visual information that greatly resembles fish. The atypical functional features of the concept *whale*, along with potentially misleading perceptual information, could increase the attraction exhibited in mouse-movement trajectories (Experiment 3 compared to Experiment 1). Finally, this prediction regarding atypical competitive trials using pictures, and the previous discussion of moderate similarity even between non-competitive categories and exemplars, suggests that substantial competition should be seen in conditions with non-competitive category alternatives when pictorial stimuli are used (Experiment 4 compared to

Experiment 2). If misleading visual features cause more equivocal information to be partially active on atypical-animal trials, even when the category label is supposedly non-competing, it should take more time for the correct category to reach maximal activation.

Participants

45 Cornell undergraduates participated in this experiment to receive extra credit for their psychology classes.

Materials and Procedure

The picture stimuli were color images of approximately equal size, collected from the Internet. Images were chosen to be easily identifiable. For example, goldfish (and whale) images were from a side-angle view. All atypical photos were whole-body images of the animals. None of the animal pictures were miscategorized persistently in Experiments 3 and 4. Besides employing pictures rather than words, the following two experiments use the same materials and procedures as the previous two.

Results

Accuracy rate for atypical animals was 96%, and for typical animals was 99% (p = .078). Once again, all incorrect trials were discarded before analysis.

Time-normalized analysis

As seen in Fig. 2.5A, the atypical-animal trials elicited substantially more curved computer-mouse trajectories than did the typical-animal trials. Collapsed across leftward and rightward trajectories, t-tests revealed atypical-typical x-coordinate differences that were significantly greater than zero (p < .05) across 61 consecutive time steps, from the 26th to 87th slices. This is considerably larger than the 38 from Experiment 1. In the repeated-measures ANOVA with binned time steps, there was a significant main effect of trial type ($F(1,44)$ = 34.5, MSe = 2012.5, p < .001), bin ($F(2,43)$ = 1773.6, MSe = 1475.3, p < .001), and interaction ($F(2,43)$ = 14.7, MSe = 1007.6, p < .001). Planned comparisons show that, in all three normalized time bins, atypical trials diverged in the x-coordinate from typical trials (p's < .05), exhibiting attraction to the competing response category.

Space-normalized analysis

By normalizing coordinates to travel from origin (0,0) to (1,1), the three real-time bins reveal a robust main effect of trial type ($F(1,43)$ = 45.6, MSe = .056, p < .001), time bin ($F(2,42)$ = 289.1, MSe = .028, p < .001), and an interaction between trial type and time bin ($F(2,42)$ = 8.1, MSe = .034, p = .001). Planned comparisons show that all three time bins (0-500, 500-1000, and 1000-1500 ms) contain significantly different mean x-coordinates for the atypical- and typical-animal trials (p's < .001).

Figure 2.5: A) Mean time-normalized mouse-movement trajectories in Experiment 3, separated for left- and right-ward responses to the correct target. Atypical picture trials (solid lines) indicate strong divergence towards the competing category. B) Space-normalized time bins show a strong attraction, and slower progress towards target for atypical trials (solid lines).

Additional measures

The difference between movement initiation latencies was significant and in the expected direction (440 ms for atypical animals vs. 383 ms for typical animals, $t(44) = 2.5$, $p < .05$). There was also a significant difference between the two conditions in the total time taken to perform the categorization (2150 ms vs. 1560 ms, $t(44) = 5.2$, $p < .001$). Movement duration differed significantly (1710 ms vs. 1177 ms, $t(44) = 7.7$, $p < .001$), and total distance traveled differed significantly as well (1017 pixels vs. 830 pixels, $t(44) = 5.4$, $p < .001$). Item analyses for all measures revealed significant differences between typical and atypical stimuli (p's $< .01$). In contrast to Experiment 1, *every* dependent measure showed a strong reliable difference between atypical and typical conditions.

Analysis of initial direction in degrees shows a significant difference between atypical and typical trials ($t(44) = 2.1$, $p < .05$), with atypical trials in fact showing a negative angle (towards competitor; -2.98) and typical trials showing a significant positive angle of 4.5 towards the correct category ($t(44) = 2.1$, $p < .05$). The negative angle of atypical trials was not significantly different from 0 ($t(44) = -.94$, $p = .35$).

Distribution analysis with the Kolmogorov-Smirnov test again shows that curvatures for atypical and typical trajectories did not have different distributions ($\chi_2 = .94$, $p = .97$). Bimodality coefficients were similar and below .555 (.503 and .457, for atypical and typical, respectively). Sample entropy measures showed much stronger effects than Experiment 1 and 2. Again using window sizes of 2-10, and .2 times the standard deviation of all x-coordinate fluctuations across trials as the tolerance, atypical trials consistently showed reliably higher sample entropy in all window sizes (p's $< .05$).

Discussion

Results indicate that the process of categorizing pictures of animals exhibited temporally dynamic spatial attraction effects in the mouse-movement trajectories. In fact, the perceptual ambiguity of the atypical animals seemed to induce greater competition effects, in both trajectory and time measures, than did the lexical stimuli, as predicted above. In Experiment 1 (with animal *words*), the spatial divergence between atypical-animal trajectories and typical-animal trajectories was significant for 38 consecutive normalized time slices, whereas in the present experiment (with animal *pictures*), the spatial divergence was significant for 61 consecutive normalized time slices. Moreover, the atypicałtypical differences in total response time, movement duration, and movement distance were also greater in the present experiment than they were in Experiment 1.

Experiment 4

Just as Experiment 1 (with animal words) had its control comparison in Experiment 2, using less-competitive alternative response options, Experiment 3 (with animal pictures) has its control comparison in the present experiment. In Experiment 4, the taxonomic class response options were the same as in Experiment 2, e.g., *mammal* and *bird* for the exemplar *whale*, but the exemplar was presented as a picture instead of a word. Thus, the alternative (incorrect) category in the atypical-animal condition here (e.g., bird) had less featural match to the exemplar (e.g., whale) than the one in Experiment 3 had (e.g., fish). However, the poor match between the salient visual properties of the picture and the correct category response (e.g., whales do not *look* much like mammals) may cause the correct category to be somewhat less competitive in the categorization process than it is when exemplars are presented as words. With closer relative competitiveness of the two categories, the alternative (incorrect) category response may be able to exert a more substantial spatial attraction effect than was seen in Experiment 2.

Participants

39 undergraduate participants signed up for this experiment for extra credit in their psychology classes.

Materials and Procedure

This experiment used the same technique and same pictures as Experiment 3, but the experimental (atypical-animal) condition involved less competitive categories. The same non-competing response options from Experiment 2 (see Table 2.1) were employed here.

Results

Accuracy rate for atypical animals was 94%, and for typical animals was 97% ($p = .58$). Once again, all incorrect trials were discarded before analysis.

Time-normalized analysis

Analysis of mean trajectories in combined left and right trajectories again showed differences in x-coordinates that were significantly greater than zero across 66 consecutive time steps, from the 28th to the 94th slices (p's $< .05$; see Fig. 2.6A). A binned repeated-measures ANOVA as in previous experiments showed a highly significant main effect of trial type ($F(1,38) = 17.0$, $MSe = 1538.6$, $p < .001$), bin ($F(2,37) = 2401.5$, $MSe = 1134.8$, $p < .001$), and an

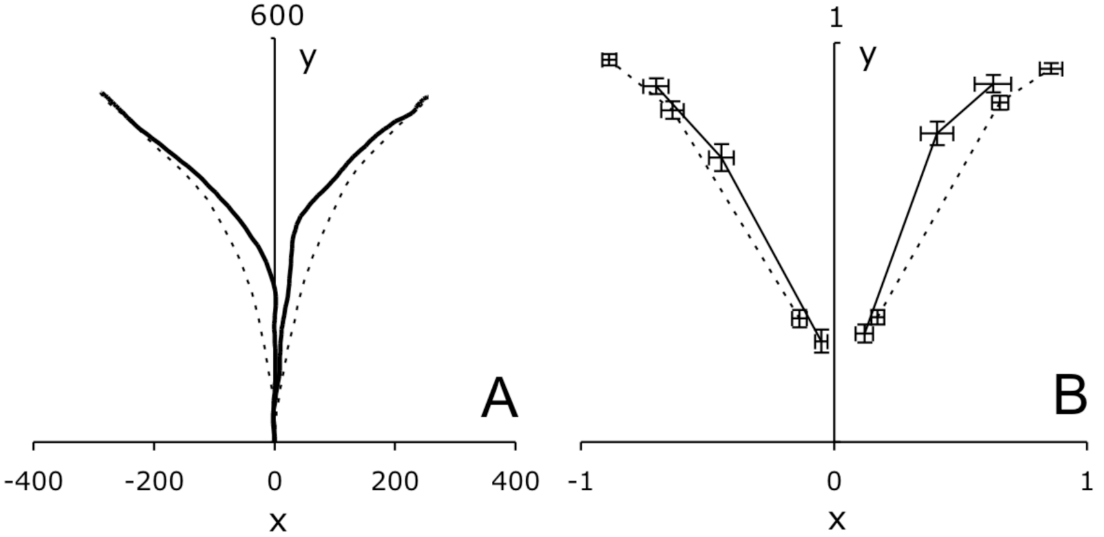

Figure 2.6: A) Mean time-normalized mouse-movement trajectories in Experiment 4. Atypical picture trials (solid lines) also reveal a significant attraction towards the competing category. B) Space-normalized time bins again show a strong but diminished attraction, and slower progress towards target for atypical trials (solid lines).

interaction with time step bin ($F(2,37) = 7.8$, $MSe = 589.6$, $p < .001$). Planned comparisons show significant differences between trial types in the second and third normalized time bins (p's < .01).

Space-normalized analysis

Fig. 2.6B shows space-normalized trajectories in the three real time bins. A repeated-measures ANOVA included a significant effect of trial type ($F(1,36) = 28.7$, $MSe = .047$, $p < .001$), time bin ($F(2,35) = 739.0$, $MSe = .016$, $p < .001$), and an interaction between trial type and bin ($F(2,35) = 9.1$, $MSe = .017$, $p = .001$). All three bins show a significant difference between trial types (p's $< .05$).

Additional measures

There was no significant difference in initiation latencies between the two condition (392 ms atypical vs. 349 ms typical, $t(38) = 1.6$, $p = .12$). As in Experiment 3, there was a significant difference between the two conditions in the total time taken to perform the categorization (1813 ms vs. 1461 ms, $t(38) = 5.2$, $p < .001$). Both movement duration (1421 ms vs. 1112 ms, $t(38) = 4.8$, $p < .001$), and total distance traveled were significantly greater in atypical trials (806 pixels vs. 747 pixels, $t(38) = 3.3$, $p < .01$). Initial angle from (0, 0) for typical trials was at 4.9, significantly greater than 0 ($t(38) = 2.1$, $p < .05$), while atypical trials were again slightly negative (-1.7), but not significant. Item-based analyses showed the same patterns of significance (p's $< .05$), though movement initiation latency showed a marginal significance ($t(17) = 8.7$, $p = .07$). Finally, a Kolmogorov-Smirnov test showed no significant difference between typical and atypical curvature distributions ($\chi_2 = 1.1$, $p = 1.0$). Interestingly, bimodality coefficients were both near but greater than .555 (.584 and .597 for atypical and typical, respectively). The value for typical trials, however, is greater than for atypical trials. Thus, if each of these distributions was genuinely the result of two independent classes of motor-movement trajectories (those that started out correct, and those that were sharply corrected in mid-flight), this bimodality cannot be attributed solely to the atypicality of the whale, seal, penguin, etc. – even the typical animal pictures elicited this pattern. Finally, just as in Experiment 2, sample entropy showed no significant differences between atypical and typical trials. As Experiments 1 and 3 showed reliably greater sample entropy for atypical trials compared to typical trials, this measure may be a critical indicator of the difference between competitive and non-competitive category trials. When both category response options are substantially competitive (Experiments 1 and 3), the mouse-movement trajectories exhibit a conspicuous complexity that may be due to the nonlinear dynamics inherent in an attractor landscape that has multiple strong attractor basins.

Discussion

Using visual images of exemplars, a non-competitive alternative category still induced considerable curvature in the mouse-movement trajectory. When divergence between the average trajectory for atypical animals and the average trajectory for typical animals is plotted over time for Experiment 3 and for Experiment 4 together (see Fig. 2.7), the competitive alternative categories in Experiment 3 show a numerically stronger attraction effect than the non-competitive alternative categories in Experiment 4. However, the two atypical-animal trajectories from Experiment 3 and 4 do not exhibit a significant difference for more than 8 consecutive normalized time steps. Nevertheless, when doing independent samples *t*-tests of the outcome-based dependent measures of the experimental trials from Experiments 3 and 4, total response time (*t*(82)=2.6, $p < .05$), total time in motion (*t*(82)=2.5, $p < .05$), and distance traveled (*t*(82)=4.8, $p < .001$) all indicated that when the incorrect alternative category was competitive (e.g,. *fish* for the *whale* picture), mouse-movement was slower and more extensive than when the incorrect alternative category was non-competitive (e.g., *bird* for the *whale* picture). Item-based

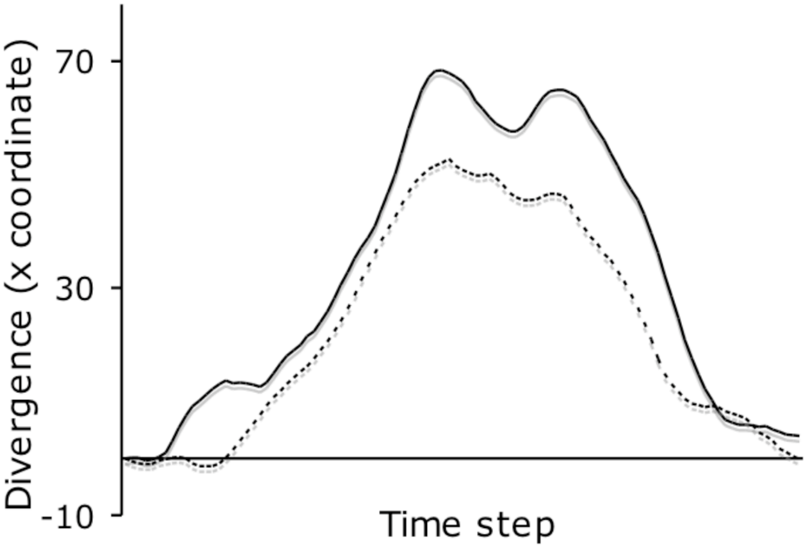

Figure 2.7: Mean total divergence (atypical trial – typical trial x-coordinates) for Experiments 3 and 4. Experiment 3 (solid line) exhibits greater divergence, though not significant for 8 time steps (see text for details).

comparisons between atypical trials between Experiments 3 and 4 also show a significant difference for distance in the expected direction (*t*(5) = 6.2, *p* < .01).

Thus, although not quite as much as in Experiment 3, images of atypical animals in Experiment 4 *did* induce computer-mouse trajectories that exhibited spatial attraction toward an incorrect category label more so than trajectories for typical animals. When words were used with this non-competitive arrangement instead of images, in Experiment 2, a much diminished trajectory curvature was observed, as was compared to its corresponding competing-label scenario in Experiment 1. As one simple explanation for this strengthened attraction in the non-competing labels in Experiment 4 (as compared to Experiment 3), the divergence between trajectories may be the result of a reduced featural match of the correct category with the visual properties of the atypical exemplar, rendering the alternative category's meager competitiveness relatively more substantial. When one then compares Experiment 4 with Experiment 3, the exceptionally long mouse movements (in time and space) for atypical-animal trials in Experiment 3 can be seen as due to that same reduced competitiveness of the correct category combined with a considerable competitiveness of the alternative category, resulting in a particularly lasting and laborious competition between the categories.

General Discussion

The lexical stimuli in Experiment 1 revealed that categories in competition lead to a dynamic spatial attraction of the resultant mouse trajectories. This competition diminishes, but does not go away, when the featurally-similar alternative category is replaced by a less-competitive alternative category in Experiment 2. Interestingly, this pattern changes when using image-based stimuli in Experiments 3 and 4. Images of atypical exemplars produce robust mouse-trajectory attraction toward both highly-competitive alternative categories and less-competitive categories alike. These results offer further insight into the nature of categorization. As mentioned in the introduction, the time course of categorization has only recently been rigorously explored (e.g., Nosofsky & Palmeri, 1997; Lamberts, 2000). Like Abrams and Balota's (1991) results with lexical decision and recall memory, these results further suggest that manual responses themselves may be reflective of a continuous, dynamic process of categorization underlying participant performance in these experiments.

Importantly, these findings may extend the information-accumulation theory of Lamberts (2000) and the exemplar random-walk model of Nosofsky and Palmeri (1997). These theories may in fact predict the above results. For example, the Nosofsky and Palmeri random-walk model involves a race among exemplars governed by their similarity to a test item. In our case, right vs. left response boxes (competing categories) attract the continuous manual movement in accord with the extent to which they fit the exemplar presented. Depending on how its representations are mapped onto motor output, this race could result in a substantial pull toward highly competitive alternative categories (as in Experiments 1 and 3), and somewhat less of a pull with less-competitive alternative categories. This basic finding that continuous, graded processing of the exemplars during categorization flows into the effectors can be seen as an extension of an iterative sequential-sampling, information-accumulation perspective. While the perspective of both Lamberts (2000) and Nosofsky and Palmeri (1997) is applied largely to the process leading up to a discrete categorization response, our results suggest that the response is itself a component of this continuous, probabilistic processing. A full synthesis may accompany future experiments that overcome some limitations of the current ones. For example, it must be acknowledged that the current research employs a small stimulus set. Future experiments may further bridge the various levels of the categorization process (from perception to response preparation and production) through more extensive stimulus sets. While the current experiments serve as an initial demonstration of these patterns using a small set of commonly known animals, there remain numerous issues to explore. These might include the resolution of item-by-item typicality, frequency, and other variables and their concomitant effects on continuous output. Despite the present limitations, these experiments demonstrate a potentially fruitful avenue for eventually mapping out an uninterrupted explanatory landscape from perceptual input to motor output.

Classical perspectives on cognitive processing as organized into discrete serial stages would likely predict that only noise would account for graded output patterns in tasks such as taxonomic categorizing – and that mean trajectories should in fact simply indicate the target response. While the current results are particularly strong for recommending against this purely serial perspective on the *process* of categorization, we would argue that they also have broader theoretical significance regarding representational issues in cognitive theories. Throughout the cognitive sciences, debate continues over the nature of conceptual representation. A central dichotomy in this debate is the format of representation, and related processing-flow issues, that underlie cognition. On one side of this debate, representations are largely characterized as discrete and symbolic, and undergo computational or algorithmic manipulation akin to a Turing machine (e.g., Fodor, 2000; Dietrich & Markman, 2003; Marcus, 2001; Pinker, 1997). On another side, there are a variety of proposals that see representational states as probabilistic, graded informational states that either undergo computational manipulation (Massaro, 1989, 1998), or are part and parcel of a system that blends content and process in its operation (Elman, Bates, Johnson, Karmiloff-Smith, Parisis, & Plunkett, 1996; Port & Van Gelder, 1995; Rumelhart & McClelland, 1986; Spivey & Dale, 2004). This dichotomy only approximates the rich spectrum of opinion in cognitive science (cf. hybrid systems such as those by Young & Lewis, 1999; Sun, 1997). Nevertheless, debate tends to focus on the extent to which each format of representation contributes most to our understanding of cognition.

A range of behavioral data is typically adduced to support one and challenge the other. A classic example of phenomena interpreted as supporting the existence of discrete-symbolic representations is the very act of categorization itself: "It follows…that if a system categorizes, then it has discrete representations" (Dietrich & Markman, 2003, p. 102). For example, early research on categorical perception of speech sounds not only fit perfectly into the traditional perspective on symbolic computation, but also led to proposals that there exist specialized processes of this sort for speech (see Massaro, 1998 for a review). Categorization is but one property of several behaviors that animals exhibit that have been proposed to support the centrality (or at least existence) of discrete-symbolic representations in cognition (for other properties, see Dietrich & Markman, 2003; Marcus, 2001). It has, however, been of central importance for understanding our cognitive architecture.

Related to recent research (e.g., McMurray et al., 2003; Nederhouser & Spivey, 2004), lexical and perceptual categorization in the present experiments revealed graded response patterns in participants' mouse trajectories (see also Spivey et al., 2005). As in Abrams and Balota (1991), it seems that the time course of categorization is reflected in the manual output from the process. In particular, given our results that competing category labels generate greater trajectory curvature, slower time measures, and longer trajectory excursions, this temporally sensitive deviation in motor output suggests partially active states underlying the time course of categorization. The upshot for cognitive representation is subtle, but important. If discrete representational states underlie cognition at some point following perceptual processing, then there must still exist some "granularity" of information, in Miller's (1982) sense, by the time this process is being converted into motor output. If this is granted, then there is a limit on the discreteness imputed to the representational states underlying categorization. Discrete representational perspectives may indeed account for these kinds of results by allowing differing levels of granularity in representational discreteness. Indeed, further pursuing the line of research in Miller (1982; see also Miller & Ulrich, 2003) may reconcile these theoretical perspectives by finding a common ground between purely continuous and purely discrete representations (cf. Dale & Spivey, in press). This would involve specifying how refined the granularity is at various stages of processing.

These studies demonstrate that echoes of continuous (non-discrete) informational states can be observed in the dynamic properties of resultant motoric responses – granularity evident even in the manual output. Any postulated discrete representational states mediating sensors and effectors must carry at least some relevant information from early graded states in order for the motor output to mimic the continuity of the sensory input. In other words, while reaction times and accuracy measures may reveal information about the decision process during discrete, algorithmic processing, the graded manual output from the system observed here indicates that even when these discrete decision processes collapse onto the effectors, there remains some granularity.

Appendix

In order to establish a basis for our criterion used in the multiple *t*-tests on time-normalized trajectories, we conducted simulations akin to the Bootstrap method (Efron & Tibshirani, 1993). The mean and standard deviation of each of 101 time steps in atypical and typical trajectories in Experiment 1 were recorded. Next, 10,000 simulated "experiments" were computed using the same mean and standard deviation. In other words, in each of these simulations, we constructed 41 model participants (*N* in Experiment 1) by constructing atypical and typical trajectories from these time-step means and standard deviations. Within each trial type, each time step (out of 101) was sampled from a normal distribution with the mean and standard deviation from the time steps of the *actual* mean trajectories, thus preserving the non-independence between time steps. We then ran *t*-tests within each of the 10,000 simulations for each time step (101 time steps, and *N* = 41). Of these 10,000, we recorded the frequency with which *sequences* of significant differences occurred. Simulations revealed significantly different sequences of 6, 7, and 8 time steps with a percentage of 3%, 1%, and .5%, respectively (see Table 2.2). A sequence of 8 consecutive significant t-tests is thus produced by chance in these simulations with less than .01 probability, and was selected as a conservative measure. This simulation therefore recommends a criterion of 8 consecutive significantly different timesteps for discerning whether the interdependent time steps in our experiments have sequences that exceed what would occur by chance alone. This criterion was also used in Experiments 2-4.

Table 2.2: Frequency (in percentage) of sequences discovered in the 10,000 simulated experiments

CHAPTER THREE

On Resilience and Versatility: The Theory-Model Distinction and Connectionism

Introduction: Of Ideals and Details

Political and social ideals swirl into the world, vanish just as easily, and routinely reappear in novel forms under different historical circumstances. The way these ideals are applied can differ widely within and across generations. Whatever the success of these applications, the ideals are often resiliently held by those who adopt them, largely because the practical details of ideals manifest themselves in versatile ways.

In some respects, a scientific theoretical framework has the properties of political and social ideals. A set of theoretical ideals in any discipline may fall out of favor, only to reappear with the advent of new conceptual, mathematical, or technical tools. In some cases, these ideals may simply reappear by a kind of reinvention of the theoretical wheel. When the details of theoretical ideals are worked out, a theorist is resilient in the face of drawbacks or preliminary failures, because the details can be elaborated with considerable versatility.

In what follows, we argue that scientists generally, and cognitive scientists in particular, harbor two personalities. The *resilient theorist* is one who sticks by her theoretical ideals, encouraging the interpretation of worldly phenomena in terms of a particular set of constructs and operations among them. The *versatile modeler* uses these ideals and *models* aspects of the world after them. The versatility with which these ideals are applied encourages the theorist's resilience, and indeed the reverse is true. These personalities are close kin, and both may inhabit a single scientist. Though very similar observations were a focus of the latter-half of last-century's philosophy of science, we aim to flesh out the consequences of these personalities and apply them, as an important conceptual basis, to debates in cognitive science.[1]

In the first part of this paper, we provide intuitive and historical support for the distinction between theories and models. The distinction as drawn is not trivial, and we describe certain consequences that importantly underlie the pursuits of resilient theorists and versatile modelers.

In the second part, we show how disregard for the theory-model distinction can lead to misdirected criticism of robust cognitive scientific frameworks. We assail arguments against connectionism in an attempt to show that ignoring the distinction and its consequences leads to some unfair assessments of connectionist theory. Finally, we describe the ways in which connectionist modelers have encoded semantic representations in models, and how this illustrates the distinction, its consequences, and the way connectionism satisfies a number of conceptual constraints.

---

[1] In fact, what we are arguing has appeared in the philosophy of science as a reaction to the Received View (Bechtel, 1988) of scientific explanation and change. Many philosophers of science have noted the steadfastness of theorists, and the latitude with which the theorists may approach data (to name a few, Kuhn, 1963; Quine & Ullian, 1970; Lakatos, 1978).

The term “theory” is indeed vague. From a Greek word roughly translated “a looking at” or “viewing,” it has been used in a variety of ways since at least 1674, when it appears in Dryden’s line “Your theories are here to practice brought, As in mechanic operations wrought.” This usage of the word, as a system of ideas or methods, is now its sole import in modern English. “Theory” in this sense appears, for example, in popular scientific press when Darwinian evolution is labeled “just a theory” by creationists (or, “intelligent design” theorists), or more peculiarly, when a friend claims to have developed a “theory” about how cats exert their telepathic powers over humans. It is crucial to point out that beyond these casual and cursory notes, we will not consider the assorted colloquial uses of “theory” and “model.” Instead, we wish to compare *theories* and *models* as explanatory entities both intuitively and with respect to the history of psychology. More importantly, the goal of this section is not to offer a definition for the terms “theory” and “model,” but to elaborate the objects of concern for resilient theorists and versatile modelers, the difference between these objects, and their respective properties (e.g., Gould, 2002, urges a curious reification or “objectification” of these explanatory objects).

To begin, consider the observation, made by a number of philosophers of science (e.g., Kuhn, 1963; Lakatos, 1978), that all scientific research is conducted in the context of a particular theoretical framework -- otherwise known as the property of “theory-ladenness.” For example, for Lakatos (1978), this framework is a *research programme*, in which broad principles and problem-solving machinery are combined to engender active and resilient scientific research. This and other descriptions of underlying theoretical frameworks are not necessarily equivalent, but overlap considerably in the way they characterize motivational forces in scientific research, and are commonly supported by providing numerous examples from the history of science.

One purpose of a framework's broad principles is to guide specific research questions. There exist, and can be observed readily in the history of science, specific research projects or goals within any particular theoretical framework. The notion that these projects occur *within* a framework simply means that they incorporate the framework's principles to explore a more detailed range of phenomena. This observation is important in the history of science. Even recently, specific research pursuits may make discoveries about the details that can cause change in the general principles (as some seek for evolutionary theory, for example). Also, because details can be worked out in a variety of ways, there can be multiple projects consistent with a framework but which compete to account for overlapping phenomena (discussed further below). These considerations suggest that general principles cannot alone characterize an entire body of scientific research. Broad theoretical principles constrain research pursuits, but these pursuits seek to press the principles into service for approaching more detailed phenomena.

Such notions as a paradigm, research programme, and theoretical framework evidently form a more abstract class of explanatory entity than the specific projects they bring about. These explanatory entities, or *theories*, are a collection of broad and common principles that guide and define a research community. But such broader explanatory entities should not be confused with the detailed investigations they underlie. The more specific and detailed accounts of phenomena they motivate are explanatory *models*, designed to instantiate theoretical principles, and elucidate their involvement in the subject matter under investigation (e.g., Cartwright, 1999, discusses myth and model in scientific explanation -- akin to theory and model, discussed here).[2]

This distinction is borne out when we consider the various theories that have shaped the history of psychology. For example, Freudian psychoanalysis is characterized by a general set of theoretical principles. However, the ways in which these are instantiated and applied to specific psychological phenomena can be considerably more complex. For example, Freudian principles could be used to explain why women fiddle with their purse, or the existence of arms proliferation (in the amusing manner one would expect; Leahey, 2001). Interestingly, given Popper's famous arguments against it, psychoanalysis may offer such lax guidance that many Freudian models, mutually incompatible with one another, could be easily concocted to explain the same behavior.

Another major trend in the history of psychology, the experimental analysis of behavior (known commonly by its moniker "radical behaviorism"), included a robust and small set of theoretical principles that formed the dominant framework in North America for over 20 years. Within this framework, there exist individual models for a wide range of phenomena, such as the emergence of object words in child language (Horne & Lowe, 1996 and Hayes, Barnes-Holmes, & Roche, 2001 present competing accounts).

---

[2] We understand that this usage of the terms may appear awkward, but the distinction as expressed will serve later discussion concerning cognitive science. In this way, Freudian stories, animal learning models, and computational models are all ``models'' that instantiate theoretical ideals.

As for the cognitive sciences, for example, Newell, Shaw, and Simon (1958) articulated a set of principles for cognition long ago that are instantiated relatively faithfully by a wide range of individual explanatory models.

Even a cursory glance at the history of psychology suggests that the distinction between these explanatory entities, though perhaps itself somewhat vague upon closer examination, can be exemplified clearly by major theoretical and historical trends. In fact, within each example, there continue to exist resilient theorists who lobby for their ideals. Freudian psychoanalysis not only persists in popular culture, but continues to be applied clinically despite its apparent drawbacks. Also, the progeny of Skinner's behaviorism still conducts research in bastions throughout the US and Europe. Finally, and relevant to the focus of this paper, the symbolic approach to cognitive science emerged as a majority party after the decline of behaviorism, and is engaged in extended debate with a new set of ideals that re-emerged two decades ago: connectionism.

Consequences of the Distinction

The distinction between theory and model so far discussed may appear rather obvious, and one may wonder whether there is justification for elucidating it. In trying to convince the reader otherwise, we argue that the distinction is not a trivial semantic quibble, but that it has a number of important consequences. These consequences provide conceptual support for the resilient theorist, and justify the versatility of modeling. And as we later illustrate, recognizing this distinction and its consequences may provide a valuable conceptual basis to help clarify confusions or misgivings in theoretical debates in the cognitive sciences.

Underdetermination of the Details

The underdetermination of scientific theory by empirical data has been recognized for decades (Bechtel, 1988). But there is a kind underdetermination that goes in the opposite direction. There may exist mutually incompatible models for a single set of phenomena, yet be entirely consonant with broad theoretical principles. The distinction implies this underdetermination because theories guide research concerns, but models work out the details in considerably more versatile ways. This is particularly true of theories whose principles lack complete mathematical formalization (such as Freudian psychology; see Meehl, 1998, for a discussion).

Evidential Relationship

The underdetermination of model by theory suggests another crucial consequence of the distinction: The evidential relationship between theory and model is more complex than sometimes assumed. As we contend below, discussion in cognitive science can too easily assume that a prominent and widely applied theoretical framework is especially disadvantaged by a small set of inadequate models.

Theories lacking mathematical formalization bear a relationship to their models much as they would to the experiments conducted within them.[3]

---

[3] In this way, a computational model is much like a micro-experiment, providing a window onto the interaction of variables in the real world in terms of the theoretical principles that motivate it.

Parallel to the demise of the "critical experiment" notion in the history and philosophy of science (Kuhn, 1963; Putnam, 1973), it is also true of modeling: An inadequate model can be remedied by a suitable reorganizing of premises or auxiliary hypotheses. In short, models can be revamped and improved. According to many thinkers, this is a characteristic of *all* theories and is not a "flaw," because in some cases resilience to anomalies is a natural trait of productive scientific theories (Lakatos, 1978).

Constrained Versatility

Another consequence of the distinction is the nature of constraints on models. The obvious constraint is the set of broad theoretical principles the model is presumed to instantiate. These *theoretical constraints* provide the overarching direction for the nature of modeling questions and methodologies. However, a model is also constrained by the phenomena under study. These *veridicality constraints* govern the way in which some of a model's parameters are included, manipulated, and interpreted.[4] A variety of perspectives in the philosophy of science lend support to such a source of constraints. For example, van Fraassen (1980) argues that the role of scientific theory and modeling as a whole is to ``save the phenomena," meaning a model should capture established aspects of the observable world.

Because models are about worldly details, they are naturally constrained by experimental data. Along with broad principles motivating

---

[4] These veridicality constraints would subsume all the criteria suggested by Christiansen and Chater (2001) to evaluate connectionist models: data contact, task veridicality, and input representativeness. Each criterion involves the extent to which the model faithfully captures properties of the world.

them, they are pressed by two sets of forces. In this sense, a model is much like a mapping between theoretical principles and the experimental data. But theoretical and veridicality constraints interact along multiple dimensions, so even theory and data together underdetermine mappings or models.

In addition, these constraints suggest a way of describing the parameters incorporated into modeling. On one hand, there is a complex parametric space that underlies the methodology suggested by theoretical constraints. On the other, you have a set of parameters that satisfy the veridicality constraints: They are representative of the variables observed or inferred in behavioral data. Although this further distinction between veridical and non-veridical (not constrained by data) parameters may seem elusive or abstruse, we exemplify this discussion in the next section by considering connectionist theory and modeling.

Though we will not outline the ideals of connectionist theory conclusively, there are a few criteria that stand out in the history of the approach. We paraphrase Smolensky's recent list (2001, p. 320), as a rather lucid and historically motivated version of connectionist theoretical principles: mental representation as distributed patterns of activity, mental processes as parallel transformation of these patterns, and knowledge as interaction of innate learning rules, architectural features, and modification of connection strengths. These principles motivate and constrain individual models developed for specific phenomena. Though models do not always meet these particular criteria, these ideals provide the primary theoretical shape of connectionism, and the models it motivates.

In the next section, we consider criticisms of connectionist theory, and discuss how some of these criticisms fail to heed the distinction between connectionist theory and its models, sometimes resulting in unfounded criticism of connectionist theory itself. Following that, we present an example of progress in connectionist modeling that illustrates the nature of constraints on models.

Connectionist Theory and Models

A number of critics fail to notice the crucial distinction between resilient theorist and versatile modeler, attitudes that likely motivate most theoretically-minded cognitive scientists. Even more pronounced, their consequences can go unheeded in important theoretical debates about connectionist theory. In this section we consider just a few recent examples of this oversight.

Criticism of Connectionism

Failure to Heed the Distinction

There are a number of recent direct and indirect criticisms of connectionist theory that fail to observe the distinction between theory and model. For example, Green (1998) argues that connectionist models cannot be theories of cognition. Given our discussion above, individual connectionist models clearly *are not* theories of cognition. His more recent arguments (Green, 2001) attack the nature of parameters available to connectionist modelers. He argues that, because the parameters are interpretable neither as higher-order cognitive states nor brain states, they cannot be adequate models of cognitive processes. However, the constraints on models discussed above indicate that his criticism is misdirected. Modelers are not realistically committed to all parameters of their models. In fact, the notion of veridicality constraints suggests that most important to simulated phenomena are those parameters that represent some feature of the behavioral data that can be meaningfully manipulated or observed as a dependent variable in a simulation. Besides, the interpretability of elements of connectionist models (e.g., the function of hidden layer activity) is perfectly amenable to a variety of analyses providing clues about function (Clark, 1990; Bishop, 1995). Connectionist models are not black boxes, because with clever technical analyses, their operation is being illuminated more clearly.

Roberts and Pashler (2000) also confound theories and models. In indirect criticism of some models, they use the terms “theory” and “model” interchangeably, and do not pay regard to the two levels of explanation prominent in the history of psychology. Roberts and Pashler argue that data-fitting practices in computational or mathematical models are poorly represented and largely uninfluential in the history and philosophy of science.[5] However, it is important to point out that recognizing the distinction between theory and model indicates that bad modeling practice does not reflect on theory. Our discussion suggests that theory can motivate good modeling practices, in which veridicality constraints are met with a variety of meaningful parameters, whose manipulation is perfectly reasonable.[6]

---

[5] We should point out that Roberts and Pashler also neglect the distinction between data-fitting as bad modeling practice, and fitting-data, a property of any good scientific model.

[6] An excellent example of justified data-fitting with parameters can be seen in Dell and colleagues’ work (Dell, Lawler, Harris & Gordon, 2004).

Roberts and Pashler do offer some important methodological points about parameter manipulation to meet veridicality constraints. However, although they do not level criticism against connectionism directly, or computational models *per se*, there are recent cases in which the distinction and all three consequences are not made clear, and result in an unfair critique of connectionist theory. Marcus (2001) offers a long examination of connectionist theory and models, with the bold goal of integrating promising features of connectionism into a classical cognitive science. Despite this seemingly productive goal, he neglects the distinction between theory and model, thereby overlooking their evidential relationship, and concludes that a whole class of connectionist model is inadequate for explaining cognition. Moreover, Marcus fails to notice the unique veridicality constraints to which *each* model is subjected. Some of his attacks on individual models are based on subjecting them to tasks at which they were not designed to succeed. If some behavioral task is not included in the model's intended constraints, it is unreasonable to expect the model to perform them, let alone to conclude from a handful of such demonstrations that the class of model is inadequate altogether for any purpose.

To summarize, a number of criticisms of connectionist models fail to make the distinction between theory and model, thereby overlooking the important consequences of this distinction. This has resulted in a number of unfair appraisals of connectionist theory, and we think that the conceptual basis offered in the first section may help focus further debate. In the following section, we illustrate that connectionism satisfies aspects of this conceptual basis through the various ways in which its versatile modelers encode semantic representations.

Connectionist Semantics

Psycholinguistics is an active area of research in which connectionist theory has exerted significant impact. As a final discussion, we will consider this development with respect to semantics, considered an essential aspect of language by philosophers and psychologists for centuries. The notion that semantics can be ignored in linguistic analyses is in obsolescence. For example, even recent formal syntactic theories rely to some extent on semantic verb categorization principles for sentence structure (Lasnik, 2002). Connectionist psycholinguistic models often incorporate semantic representations to some extent. In this section, we consider the ways in which theoretical and veridicality constraints have impinged upon these representations.

We illustrate here that connectionist models using semantic representations range from drastic idealizations faithful to theoretical constraints, to scaled-up systems permitting the satisfaction of more complex veridicality constraints. This would satisfy what Steedman (2001) calls a "project for a scientific psycholinguistics," in which the explanation of learning language would start earlier in initial perceptuo-motor learning. In fact, semantics, he claims, is grasped so poorly because the primitives of language are "grounded in very obscure ways in our physical, social, and intellectual interactions with the real world." (p. 364) Connectionist models are vastly improving their approach to these veridicality constraints, while still maintaining connectionist ideals. We first discuss some clear idealizations, then move on to the improvements.

A very simple example of encoding lexical semantics is offered by Oliphant (2002). In his simulation, a population of Hebbian networks reconcile an atomic semantic meaning with a set of available signals. Indeed, this is even an idealization upon the principle of distributed representations described by Smolensky earlier. However, this idealization permits Oliphant to speculate on the emerging learnability of communication systems, rather than their becoming more innately circumscribed. But this approach by no means approaches Steedman's semantic desiderata.

Another method of encoding lexical semantics is by groups of features. Plaut and Gonnerman (2000) illustrate this by using randomly generated features. Their model is an excellent application of this idealization: Their representational system is well-controlled and can effectively test the products of different kinds of morphological systems. Other modelers handpick their features (e.g., Hinton & Shallice, 1991). This method permits control over lexical semantics, but lacks further veridicality: The features are at the whim of the modelers' semantic intuitions. An improvement over this approach is achieved by McRae, de Sa, and Seidenberg (1997), in which they canvas hundreds of study participants for features of lexical items. The result is a list of 9,618 features composing unique semantic vectors for 190 words. This approach permits modelers to use the intuition of native speakers to generate feature lists, but fails to approach the specific perceptuo-motor issues Steedman addresses.

Most recently, Plaut (2002) has offered a promising model that incorporates many of the features of Steedman's project, meets the broad theoretical ideals of connectionism, and involves expansive satisfaction of veridicality constraints. In his model, semantic representations, rather than being amodal or entirely hand-coded, emerge as the consequence of interacting sensory information that converges on a cross-modal layer of units. The network has a total of 28,000 connections, primarily in the cross-modal ``semantic" layer. The integration of sensory information at the semantic layer obeys a kind of sensory topography: Different regions of the semantic nodes are devoted more or less strongly to one specific source of sensory information. Tactile and visual information were processed by this network while it learned to associate a lexical item (encoding in terms of phonological features) or a particular action (gesturing). Even at the level of detail described here, it is clear that this model advances the state of the art considerably by moving beyond previous attempts at encoding semantics and taking seriously the kind of constraints Steedman suggests.

In summary, connectionist modeling illustrates well the notion of competing theoretical and veridicality constraints. Existence proofs, much like Oliphant's above, can be seen as models whose purpose is to loosely satisfy veridicality constraints, but demonstrate the success of theoretical constraints and ideals to capture a broad pattern of data. More substantive models, such as Plaut's, aim to apply theoretical ideals to difficult details (in his case optic aphasia), and in due course, satisfy veridicality constraints more broadly than any competing theory's models have yet offered (e.g., modular-computational theory of mind; Fodor, 1983, 2000).

Conclusion

We have presented a conceptual basis for considering theories and their instantiations, and exemplified them through connectionism. The consequences of this conceptual basis have been ignored by some critics, and they sometimes level unfounded criticism against connectionist theory. Despite this, connectionism exhibits all the hallmarks of a good scientific framework. Its theorists are resilient, and its methods allow for versatile modeling. It meets both broad and explanatory theoretical constraints, while actual practice within the approach continually scales-up to meet complex veridicality constraints.

It is important, however, to remark that, as in any scientific discipline, there is always progress to be made. Not all areas of connectionist psycholinguistics, or other areas in which models are adopted, involve this important interplay of existence proofs and substantive modeling, through which theoretical and veridicality constraints are satisfied. It is crucial always to remain vigilant of limitations on the ideals and details of any scientific research, lest we overlook ways to improve its ideals or learn how to better apply them (a mistake made recently by Marcus, 2001, who criticizes connectionism, but omits any reflection on the limitations of opposing approaches to which he subscribes). We only hope that the proposed conceptual basis may encourage modelers, connectionist or otherwise, towards this goal.

The next chapter explicitly seeks a model that satisfies the theoretical constraints on the perspective discussed in the introductory chapter: Cognitive processes flow into action continuously. The next section presents a simple connectionist model that captures basic aspects of categorization, along with exhibiting the important theoretical constraint of continuous processing through time.

CHAPTER FOUR

The Relationship Between Decision and Action: Simulating Dynamic Properties of Responses in Categorization

Introduction

Cognition is often thought to happen in places between the sensors and effectors, with sharp delineations at these boundaries (e.g., Fodor, 1983; Pylyshyn, 2000; see Rosenbaum, 2005, for relevant discussion on motor control). This common account envisages cognition collapse its decisions onto the effectors, directing various motor systems in an all-or-none fashion. For example, numerous theories of attention posit central processing separate from systems responsible for guiding action (e.g., Posner, 1980; Posner & Petersen, 1990). In rapid, frequent, and metabolically cheap motor output, such as the ballistic saccades of the eyes, this perspective seems most apt. When the eyes are drawn by an external stimulus, such as a sound source, the processing that has led to the saccade is programmed over a couple hundred milliseconds, and once it reaches premotor regions responsible for saccadic movement, fires off a motor instruction that may differ from moment to moment only because of noise. In other words, once the command has been issued, and the eyes instructed on their movement, a ballistic, linear motion to the target should be observed (e.g., Becker, 1991).

Recently, a growing body of research has revealed that motor systems in fact interact more richly with cognitive processes. Even in the case of saccadic movements, Doyle and Walker (2001) demonstrate that these rapid eye movements may reflect underlying attentional processing of visual cues.

They show that the eye-movement trajectory in a saccade to a given location actually *curves* depending on distractor or cue stimuli (see also Sheliga, Riggio, & Rizzolatti, 1994; Gold & Shadlen, 2000; Gaveau, Martin, Prablanc, Pelisson, Urquizar, & Desmurget, 2003). Rather than a direct, linear motion, the eyes reflect underlying processing in the graded nature even of saccades. This occurs *after* the cognitive processing has given way to them, and may thus be described as a "post-decision" dynamic property of a saccade trajectory, where the purported "decision" event is defined as the point at which motor movement is initiated. Moreover, when multiple fixations of the eye are tracked during an experimental task, they reveal what may be described as a dynamically changing probability distribution over graded representational states (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; McMurray, Tanenhaus, Aslin, & Spivey, 2003). In both individual saccades and semi-continuous tracking of multiple eye-movements, dynamic properties of motor output reflect underlying cognitive processing.

Perhaps more compelling is evidence that *manual* responses exhibit a range of dynamic post-decision characteristics that reveal processing. For example, force and velocity of responses after initiation vary concomitantly with frequency in a lexical decision task (Abrams & Balota, 1991; Balota & Abrams, 1995), and response and stimulus probability in simple reaction-time tasks (Mattes, Ulrich, & Miller, 2002; Ulrich, Mattes, & Miller, 1999; see also Balota, Boland, & Shields, 1989; Osman, Kornblum, & Meyer, 1986; Tipper, Howard, & Jackson, 1997). More recently, Spivey, Grosjean, and Knoblich (2005) used computer-mouse trajectories to study spoken word recognition. These trajectories show continuous dynamic activation of multiple competing representations during spoken word recognition (see also Allopenna, et al., 1998, for related eye-movement results). In a similar series of experiments, Dale, Kehoe, and Spivey (in press) show that computer-mouse trajectories indicate competition between categories when classifying atypical animal exemplars. The dynamic characteristics of the response, in the form of mouse trajectories, revealed an attraction towards a featurally-similar category label (e.g., *fish*) for atypical exemplars (e.g., *whale*) compared to control trials (e.g., *cat*).

Numerous studies have suggested that properties of motor output itself shed light on their originating cognitive processes. Perhaps more importantly, these studies show that processing flows in systematic ways into motor behavior, rather than simply collapsing onto them to generate a categorical response. They may indeed recommend a “cascadic flow” perspective on cognition that sees information flow continuously from sensors to effectors (McClelland, 1979; Balota & Abrams, 1995; Spivey et al., 2005), permitting the emergence of these post-decision response dynamics. This perspective changes the way in which we understand how cognition becomes action, their relationship, and possibly their interaction.

While a simplistic perspective on their relationship -- envisioning cognition collapsing onto action -- may have motivated some to neglect the important psychological relevance of action (Rosenbaum, 2005), these discoveries of dynamic response characteristics suggest numerous lines of inquiry. One such question, and the focus of the present paper, is the following: What kind of interaction between cognition and action gives way to dynamic post-decision characteristics of responses?

In this paper, we employ a neural network model of categorization to explore this question. We aim to further elaborate the possible relationship and interaction between cognition and action using a localist attractor network that categorizes animal exemplars. While the model is simple, it provides an explicit terrain for exploring dynamic post-decision response patterns, and what they might tell us about the cognitive system.

In what follows, we briefly review research on categorization and the role typicality plays in common accounts of categorization. We then present two neural network simulations. The chosen network architecture permits comparison of different relationships between internal processing of animal exemplars, and an output component that models an explicit motor response. In the first simulation, we explore this by investigating a parameter space that specifies the network's decision-action relationship. Network conditions that vary this relationship are compared in how well they fit with human experiments that show dynamic characteristics of motor responses that illuminate cognitive processing. In the second simulation, we generate a prediction about how response dynamics should reflect finer category structure: Graded typicality results in graded response dynamics.

Categorization and Typicality

Theories of categorization have grown in sophisticated ways over several decades, from classical set theoretic approaches, to rich knowledge-based accounts of categorization (see Medin, 1989, and Murphy, 2002 for reviews). In the current study, we make use of the graded nature of category structure: Categories, whatever their origin, have members that lie along typicality gradients. For example, an animal can be more or less typical as a member of an animal category, such as of mammals. The members of the mammal category thus reveal a typicality gradient, with cats being typical, and whales being considerably less typical. This results in a variety of experimental effects. Typical members are recognized faster (Rips, Shoben, & Smith, 1973), more consistently (McCloskey & Glucksberg, 1978), have many features in common (Rosch & Mervis, 1975), and can facilitate language comprehension (Garrod & Sanford, 1977).

Dale et al. (in press) make use of this property of category structure to uncover post-decision response dynamics in human participants. Computer-mouse trajectories were recorded during a simple categorization task. Participants categorized an animal exemplar by clicking the mouse on one of two category choices. Mouse-movement trajectories consisted of a movement from the center-bottom of the screen, to the correct target on the left or right (beside which was a competing category label). Target trials used atypical animals (e.g., *whale*) and included an incorrect competitor category that had considerable overlap in terms of semantic and visual features (e.g., *fish*). Though participants responded by clicking the appropriate category (e.g., *mammal*), mouse-movement trajectories exhibited substantial attraction toward the competitor category. Competing activation of the incorrect category in these trials was evident even in the properties of the resultant motor output, and not simply in the decision processes leading up to it.

Categorization may be a particularly good cognitive process to continue exploring these properties of motor output. Numerous studies have employed fairly simple experimental methodologies, such as cued fixation (e.g., Doyle & Walker, 2001) or cued reaching (e.g., Tipper et al., 1997), and traditional reaction-time tasks (e.g., Ulrich, Mattes, & Miller, 1999). Categorization is a relatively “higher-order” process, involving more processing of exemplars prior to the manual response. In addition, while lexical decision tasks also serve as good indicators (e.g., Balota & Abrams, 1995; Spivey et al., 2005), research on category structure avails itself of a broad range of stimuli, both of real and artificial exemplars, permitting a wide range of parameters to be explored (e.g., Murphy, 1991). More importantly for the present paper, theories of category structure have often made use of feature-based explanations. With very few exceptions, theories of categorization have proposed evaluation of object features, and in some manner comparing these features to prototypes or exemplars in memory -- perhaps in the context of a knowledge domain (see Medin, 1989). This aspect of categorization theories makes this cognitive process amenable to constraint-based neural network architectures. In the following section, we present a simulation that processes features of categories in parallel, and can model the time course of categorization of typical and atypical animal exemplars.

Before setting forth the current simulations, it is important to note that there exist a number of quantitative models for categorization, including basic retrieval and recognition mechanisms (e.g., Nosofsky, 1987), category learning and organization (e.g., Love, Medin, & Gureckis, 2004), and even typicality (e.g., Hirahara & Nagano, 2003; Smith, 2002) and the time course of categorization (Nosofsky & Palmeri, 1997; Lamberts, 2000). These and related models have been applied in many domains relevant to concepts and categorization (e.g., Collins & Loftus, 1975; McRae, de Sa, & Seidenberg, 1997; Medin & Shaffer, 1984; Nosofsky, 1987, 1988, 1991; Rehder & Murphy, 2003; McClelland & Rogers, 2003). Despite such a broad range of models and their application, none makes a distinction between the decision mechanism and a component that implements the overt response. For example, Nosofsky and

Palmeri (1997) and Lamberts (2000) offer prominent models for the time course of the *decision* leading up to the response, but not the dynamics of the response itself. As described above, the dynamics of motor output has a character that systematically reflects underlying cognitive processing. We make use of a relatively transparent neural network model that facilitates the inclusion of a motor-response component -- the response dynamics of the model discussed below is a natural extension of the architecture of the system itself. The model therefore includes decision and response components that are very similar in their operation, and the network lends itself to a straightforward manipulation of parameters to explore the relationship between the categorization process, and the response dynamics it generates.

A second important point about the model is that it finesses some debate about the nature of categorization. In particular, exemplar-based and prototype-based theories of categorization have been attacked and defended with little abatement for the past 20 years (e.g., recently, Smith & Minda, 2000; Nosofsky & Zaki, 2002). We do not address these broad debates, but for simplification assume a basic prototype for each animal category used: Mammals, fish, etc., are assumed to have ideal, familiar, frequent, etc. features that define a prototypical exemplar. While this simplifies the presentation of the model here, it would not be difficult to implement a learning mechanism in the model in which certain features are weighted by exposure to exemplars, thus influencing the model by specific category exemplars. Moreover, the way in which we integrate a response component in the model may be extended to the models discussed in the previous paragraph. We discuss a few such examples following the simulations, in general discussion.

Simulation 1

Network Architecture

We use normalized-recurrence to simulate the time course of categorization. This localist attractor architecture has been used to model a range of cognitive processes, including phoneme perception (Spivey & Dale, 2004), spoken-word recognition (Spivey et al., 2005), online sentence processing (Spivey & Tanenhaus, 1998; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, & Hanna, 2000), and in modeling the time course of visual search (Spivey & Dale, 2004; Spivey-Knowlton, 1996). The architecture is *localist* because individual units stand for specific features of the animal exemplars. For example, when having the network categorize *cat*, one unit in the network might be active to represent the feature representing the animal's habitat, <*land*>. The architecture is an *attractor* network because iterated updates of its unit activations lead the network towards a stable state. By applying a set of activation-update rules, the activations of the network's units are expected to asymptote on particular values, usually with one unit obtaining maximal activation, and competing units approaching minimal activation.

The network functions according to two basic computational principles. First, multiple feature units simultaneously constrain the network's behavior. Second, this parallel feature processing is integrated in a layer of units representing alternative outcomes of the model, in which one unit achieves maximal activation over time. Fig. 4.1 presents the current model. It consists of an array of feature layers, in which each unit represents a particular property

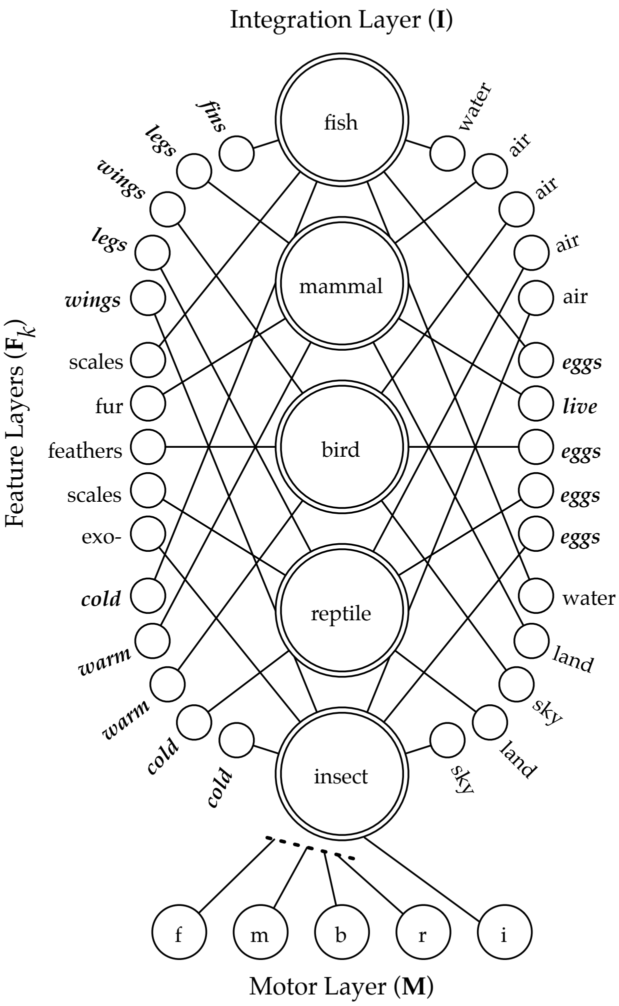

Figure 4.1: Visual depiction of the model, with interconnected layers. The motor layer (M) has one-to-one connections from integration nodes (e.g., reptile to "r"; and mammal to "m").

of animal exemplars. These layers feed into an integration layer, in which the outcome of categorization is assessed by iterated parallel processing of the constraints (see Table 4.1).

To further illustrate how the model captures the time course of categorization, we can present an exemplar to the model by setting the appropriate features in the layers to an activation of 1, and all those not relevant to 0. For example, when setting features representing the exemplar *whale*, the values for LIMBS features would have 1 for <*fins*>, and 0 for all other units. The activity of all these layers serves as input at the integration layer by taking a sum over all relevant nodes. For the processing of *whale* features, the net input to the *mammal* node would be the sum of the activation levels of the units representing mammalian features – in this case, 0 for <*legs*>, but 1 for <*air*> as source of oxygen, and so on. When processing *whale*, the *fish* node will thus receive some net input from the <*fins*> feature unit.

The integration layer is then updated by normalization: The units are made to sum to 1 all together by dividing these net input values by the total net input to the integration layer across all units. This normalized activation then feeds back into the feature layers, these are then normalized, and the cycle is continued until the integration layer becomes stable. Fig. 4.2 represents such a simulation run for *whale*, in which repeated iteration results in stable and maximal activation of the unit for *mammal* in the integration layer.[1]

Formal presentation of this models' functioning is straightforward. Net activation into the *i*-th integration unit is the total sum of its relevant feature units across all the feature layers (the *i*-th unit in each layer). From this net input, the output from this integration unit is obtained by dividing this input

---

[1] In this example and all simulations, to increase the "task veridicality" (Chriatiansen & Chater, 2001), and make similar its trials to those of human participants, the model was permitted to "see" the category labels for 5 iterations prior to having the exemplar presented. This served to lend some small activation to the features relevant to these potential categories.

Table 4.1: Categories and their prototypical features used in the model

Feature	Category Fish	 Mammal	 Bird	 Reptile	 Insect
OXYGEN	water	air	air	air	air
BIRTHING	eggs	live	eggs	eggs	eggs
HABITAT	water	land	sky	land	sky
BLOOD	cold	warm	warm	cold	cold
SKIN	scales	fur	feathers	scales	exo-
LIMBS	fins	legs	wings	legs	wings

activation by the total sum of activation into all integration units. With $\mathbf{F}_k$ representing the k-th feature layer in the set of layers $\mathbf{F}$, and $\mathbf{I}$ the integration layer,

$$net_{\mathbf{I},i,t} = \sum_{\mathbf{F}}^{k} a_{\mathbf{F}_k,i,t-1}$$

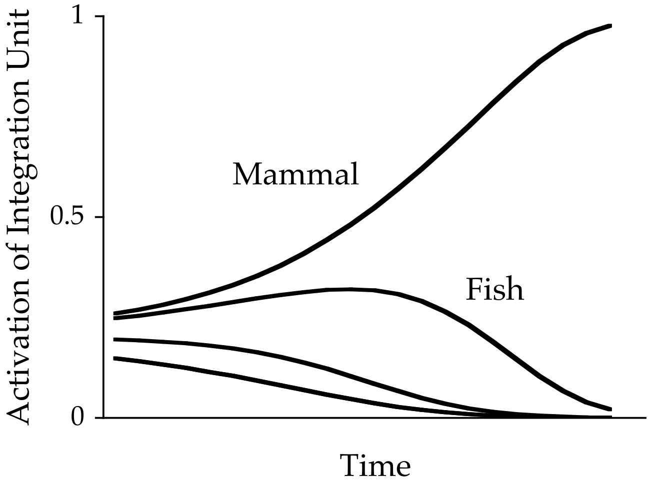

Figure 4.2: An example run of the network over multiple iterations for the exemplar *whale*. At first, categories (integration unit activations) are relatively similar, and gradually the unit representing mammal gains maximal activation, and incorrect categories approach 0. Due to featural overlap between whale and the category fish, this unit more successfully competes with mammal than the others.

$$a_{\mathbf{I},i,t} = net_{\mathbf{I},i,t} \,/ \sum_{\mathbf{I}}^{J} net_{\mathbf{I},j,t}$$

where $net_{.,i,t}$ and $a_{.,i,t}$ represent net input and activation of the *i*-th unit of the given layer at time step *t*. The value $net_{\mathbf{I},2,t}$ is the net input to the second unit of the integration layer, receiving input from the second feature unit of all 6 layers of **F**. The subsequent activation of a feature unit is determined by the sum of its previous activation and that activation multiplied by the corresponding input from the integration layer. Each feature layer then outputs a normalized activation, as in the integration layer. This process continues until one unit in the integration layer (or motor layer) reaches criterion activation (see below).

$$net_{\mathbf{F}_k,i,t} = a_{\mathbf{F}_k,i,t-1} + a_{\mathbf{I},i,t} \bullet a_{\mathbf{F}_k,i,t-1}$$

$$a_{\mathbf{F}_k,i,t} = net_{\mathbf{F}_k,i,t} / \sum_{\mathbf{F}_k}^{j} net_{\mathbf{F}_k,j,t}$$

There are a number of benefits to this simple architecture. First, its inner-workings are directly scrutable. While it is important to seek scaled-up systems that fit data from categorization (e.g., Love et al., 2004) or other processes (e.g., language processes, Rohde, 2002; Plaut, 2002), simple models that capture core theoretical principles may serve as explicit and transparent accounts for basic patterns of behavioral or neurophysiological data. One such core theoretical principle, a second benefit of this model, is conceiving cognitive processes as subject to simultaneous informational constraints. This constraint-based approach to the time-course of categorization fits with perspectives on related cognitive phenomena (e.g., Simon & Holyoak, 2002), particularly psycholinguistic processes (e.g., Bates & MacWhinney, 1989; MacDonald, 1994; Seidenberg & MacDonald, 1999), and other feature-based approaches to semantic knowledge (see McRae, 2004, for a review). Finally, this transparency and theoretical property are implemented in a simple system that operates through iterated updates of its activations. This makes normalized-recurrence particularly suitable for capturing temporal properties of cognition, an important goal for models of a cognitive process such as categorization (Nosofsky & Palmeri, 1997).

As shown above, this model permits tracking of competing category nodes, and exhibits a time course showing one winning over the others. In order to map this categorization model onto a simulated response, we supplement it with an additional bank of nodes that receives input from the integration layer. This "motor" bank of units, denoted **M**, represents the outcome response, and also exhibits a time course over iterations of the model (see Fig. 4.1). Just as the integration layer approaches a stable decision through input from feature layers, the motor layer does so through input from the integration layer. At a given time t_α, some number of time steps over which integration and feature layers have interacted, these **M** units begin to receive activation from the integration layer in a manner similar to how the feature layers receive integration activation.

$$net_{\mathbf{M},i,t} = a_{\mathbf{M},i,t-1} + a_{\mathbf{I},i,t_\alpha+t-1} \bullet a_{\mathbf{M},i,t}$$

The activation of the i-th unit $a_{\mathbf{M},i,t}$ is then similarly normalized. The parameter t_α may be varied, allowing **M** to receive input from the categorization decision at different points in time. We choose to vary this parameter by observing the "confidence" of the categorization decision -- the maximal activation found in the integration layer, denoted here as α. The integration and feature layers may therefore be permitted to interact for a period of time (up to iteration t_α) before **M** receives input from the integration layer when the maximal activation in the integration layer is α or higher.

As outlined further below, this allows a number of parameters specifying the interaction between **I** and **M** to be modified. First, a "collapsing" of the integration layer's information before it outputs to **M** can be implemented. This has the effect of producing a serial transfer of information from the categorization process into the generation of a response. Second, the point in time at which **M** begins to receive input from **I**, and begins to update its activation towards a stable output response, can be modified. Finally, **M** may be included in the set of layers **F** that feed into the integration layer. This allows the process of categorization to be influenced by available responses in the task.

Procedure

In the following simulation, we compared three different initial network conditions. An *atypical-competitive* condition involved initializing a non-prototypical exemplar (e.g., *whale*), and a featurally-similar competing response possibility (e.g., *fish*). An *atypical-noncompetitive* condition was different only in activating a less saliently competing response unit (e.g., *bird*). Finally, in a *typical* condition, (proto)typical feature values were activated, with a randomly selected competing response option. Table 4.2 presents the basic outline of simulation runs. Each trial in the conditions was selected to have some overlap with Dale et al. (in press), in which human participants were subjected to similar trial types.

The beginning of a simulation run involved setting feature layer units to their relevant values, and turning the two relevant response units to .5 to have equibiased initial response options. According to the equations above,

Table 4.2: Initial network conditions: Exemplars and response options

Condition	Exemplar	Correct / Incorrect category
atyp-competitive	whale	mammal / fish
	penguin	bird / fish
	bat	mammal / bird
atyp-noncompetitive	whale	mammał /reptile, bird, insect
	penguin	bird / reptile, bird, insect
	bat	mammał / reptile, fish, insect
typical	prototype	mammal / (4 others)
	prototype	bird / (4 others)

activation then feeds into the integration layer, and back, until the motor layer reached criterion activation (.95).

As mentioned above, the parameters of this attractor network may be modified so as to vary the relationship between the categorization units in **I** and the output units **M**. Three separate sets of runs using these network conditions were conducted to explore this relationship. In the first, we merely demonstrate that the model captures the intuition regarding serial output to motor processes: Categorical output to motor units produces a categorical time course of responding. In a second and third set of runs we reveal that the graded response dynamics exhibited in the time course of **M** is influenced by two factors. First, how much pre-decision processing occurs before **I** comes to generate a response in the motor layer. And secondly, how interaction between motor and integration layers contributes to graded effects observed in human experiments. Each of these conditions is outlined individually below.

Collapsing Information Before Output. This condition is a simple demonstration of the absence of graded patterns the model's "motor output." This occurs when the integration layer **I** collapses its information once a criterion activation has been reached. For this simulation, we choose a threshold of .5, while any value will work. Once $a_{\mathbf{I},i,t}$ for some unit i achieves a value of .5 (and all others will thus be .5 or less), $a_{\mathbf{I},i,t}$ is set to 1, and all other categories to 0. This loses the graded information contained in **I**, and then propagates to **M** to generate a response. The parameter t_α is thus set to that particular iteration's time step when there is .5-activation in one unit of the integration layer ($\alpha = .5$).

Varying **α.** The point in time at which activation flows from **I** to **M** can also be varied. This can be accomplished by varying the confidence threshold (α) at which the integration layer begins to influence the motor layer. As mentioned, t_α is defined as the point at which one integration layer unit has activation of α or greater in each subsequent iteration:

$$t_\alpha = t' \text{,} \quad \textit{where} \; \max(a_{\mathrm{I},t'}) \geq \alpha \;\; \textit{and} \;\; \max(a_{\mathrm{I},t'-\varepsilon}) < \alpha, \; \varepsilon = 1, \; ..., \; t' - 1$$

We do this while maintaining the non-integral activation values in layer **I**. We chose three levels of threshold α that span a reasonable range of possible activation of the integration units: 0, .4, and .8. This parameter will reveal in the model the amount of processing that may lead to or diminish dynamic post-decision dynamics in the activation of **M**. We hold the interaction parameter (β; see next section) constant at 1 while modifying t_α.

Varying Interaction between **I** *and* **M**. Finally, we examine whether interaction between integration and motor layers contributes to graded output in the model, and as observed in human experiments. To do this, we add a term to the net input to unit *i* in **I**:

$$net_{\mathbf{I},i,t} = \sum_{\mathbf{F}}^{k} a_{\mathbf{F}_k,i,t-1} + \beta \bullet a_{\mathbf{M},i,t-1}$$

This permits activation in **M**, set at the beginning of the simulation run, to impinge on the time course of the categorization decision -- activation in **I**. The interaction parameter β permits variable interaction from **M** to **I**, and we use a broad range of values again: 0, .5, and 1. We hold t_α constant across these parameters using a threshold of α = 0.

Results

Collapsing Information Before Output. Fig. 4.3 presents mean motor activation of networks when information is collapsed for the winning category unit (e.g., *mammal* for *whale*). All run conditions were normalized into 10 time steps so they could be overlaid (Spivey et al., 2005; Dale et al., in press). The figure presents the iteration after competing response units are set at .5, showing the subsequent 9 normalized time steps before the motor units reach criterion of .95 or greater. All conditions, whether atypical or typical, exhibit the same time course of output. This perhaps trivial result merely serves to illustrate that adopting the all-or-none assumption in the current architecture naturally leads to identical motor responses once the categorization decision has been established.

Varying α. Fig. 4.4 presents the effect of increasing or decreasing the threshold at which **I** begins to feed into **M**. The normalized time course of correct category units becomes more similar across conditions as threshold increases. The earlier the integration layer feeds into motor, the greater the competition effect seen in the atypical-competitive activation change. As an additional check of this pattern, we ran two further parameter values having t_α established at integration thresholds 0, .2, .4, .6, and .8. The average difference between atypical-competitive and typical trajectories in the three middle time steps diminishes significantly as this parameter is increased ($r = -.97$, $p < .005$).

Varying Interaction between **I** *and* **M**. Fig. 4.5 shows that feedback from **M** into **I** is not required to generate the graded patterns in the atypical-competitive condition. However, increasing feedback from **M** generates a

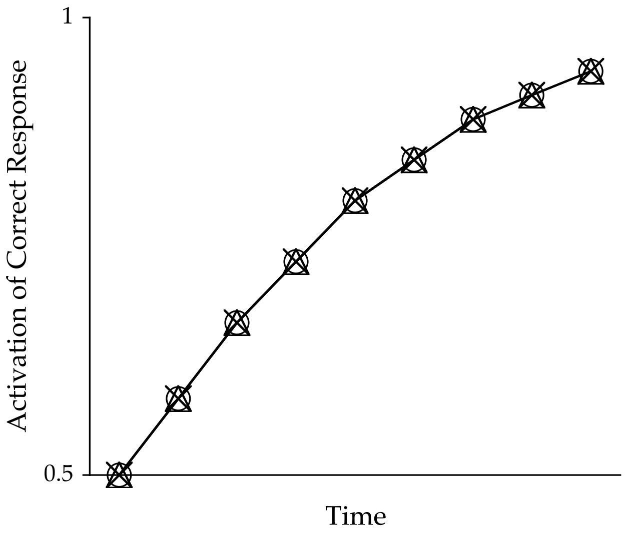

Figure 4.3: No matter the initial condition of the network, the time course of the motor response is precisely the same (cross = atypical-competitive; triangle = atypical-noncompetitive; circle = typical).

slight competition in the atypical-noncompetitive runs. Again, we ran another two parameter values (β = .25 and .75), and tracked the average difference between the middle three time steps for atypical-noncompetitive and typical conditions. This difference increases significantly with β ($r = .98$, $p < .005$). In other words, there emerges a disparity in the output dynamics of typical and atypical-noncompetitive runs when one allows motor and integration layers to more fluidly interact.

Mapping to Human Data. To investigate the relationship between the simulation and human data, we analyzed x-coordinates in manual trajectories drawn from Dale et al. (in press). A number of measures may be used, such as y-coordinate and Euclidean distance to target. However, the x-coordinate more closely represents proximity to correct or incorrect target (since, in the

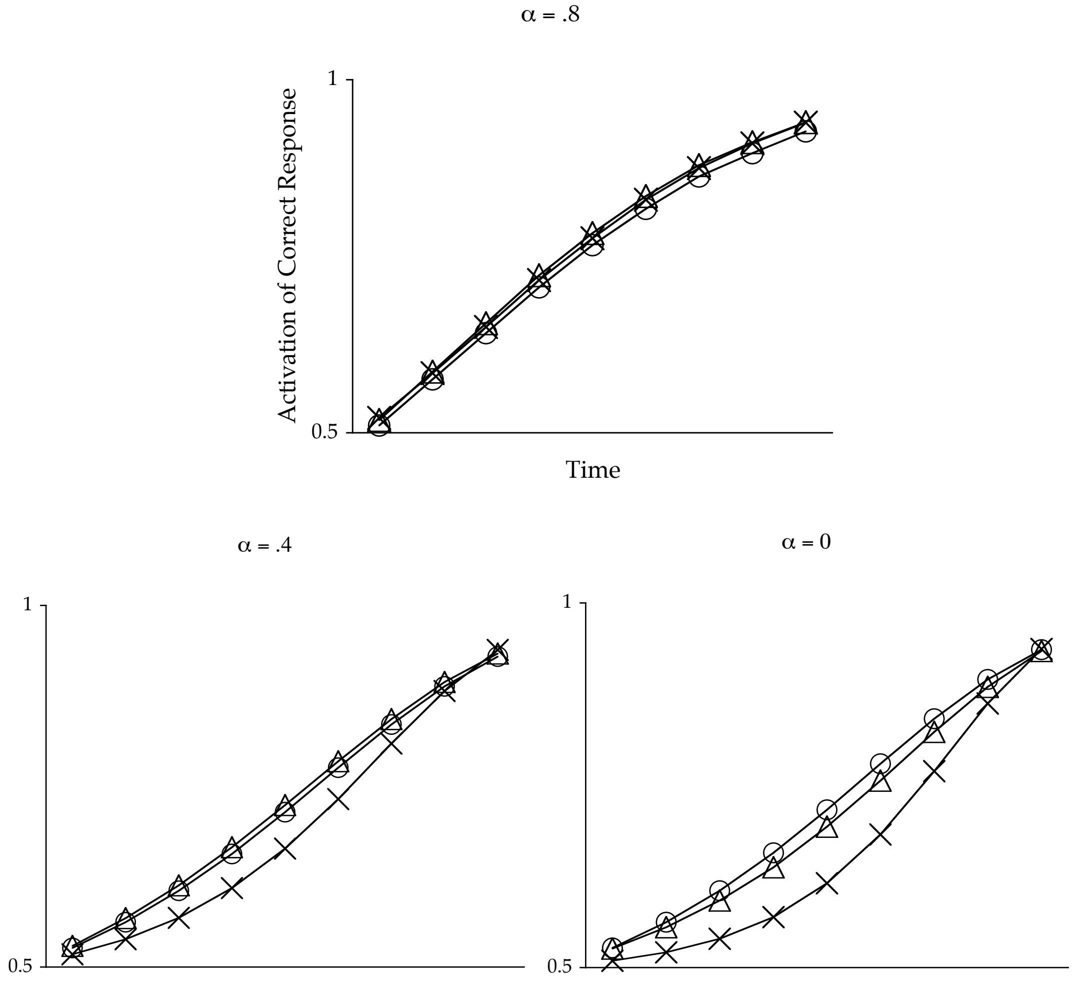

Figure 4.4: As the confidence threshold decreases, differences between conditions become more marked. The sooner the integration layer inputs to response units, the more extensive the effect of typicality (cross = atypical-competitive; triangle = atypical-noncompetitive; circle = typical).

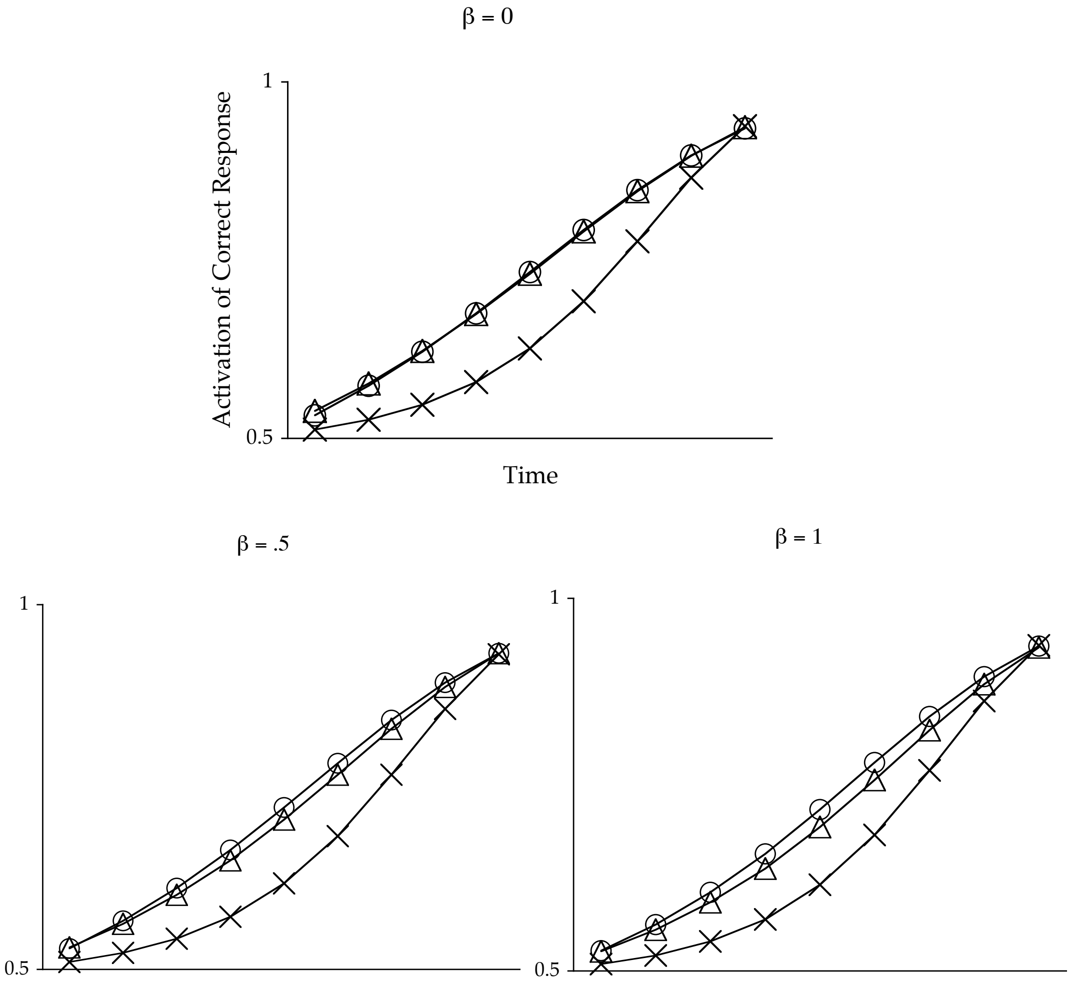

Figure 4.5: As the interaction parameter for **M** and **I** increases, differences between typical and atypical-noncompetitive become more marked. The more the integration layer receives feedback from available response options, the more extensive the effect of typicality even if the response options are not obviously competing (cross = atypical-competitive; triangle = atypical-noncompetitive; circle = typical).

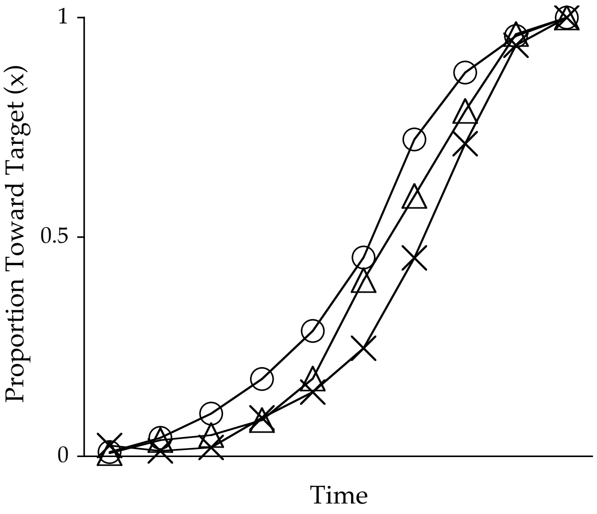

Figure 4.6: Normalized x-coordinate location, giving proportion of x-coordinate within a trial completed, grouped into the same number of time bins used in the simulation (cross = atypical-competitive; triangle = atypical-noncompetitive; circle = typical).

experiment, y-coordinates are the same for each category response label). In addition, differences in x-coordinates have served as the basis for comparing response dynamics in previous work (Spivey et al., 2005; Dale et al., in press).

Data from the lexical categorization tasks in Dale et al. (in press) are shown in Fig. 4.6. This graph depicts similarly normalized time courses of the same trial types in human experiments as the manual response's x-coordinate reaches its target. The figure bears resemblance to network conditions that involve interaction between integration and motor layers in both directions. In fact, in both perceptual (i.e., pictures) and lexical categorization, Dale et al. (in press) observed curved motor trajectories in trials akin to atypical-noncompetitive network conditions here, along with the more intuitive effect

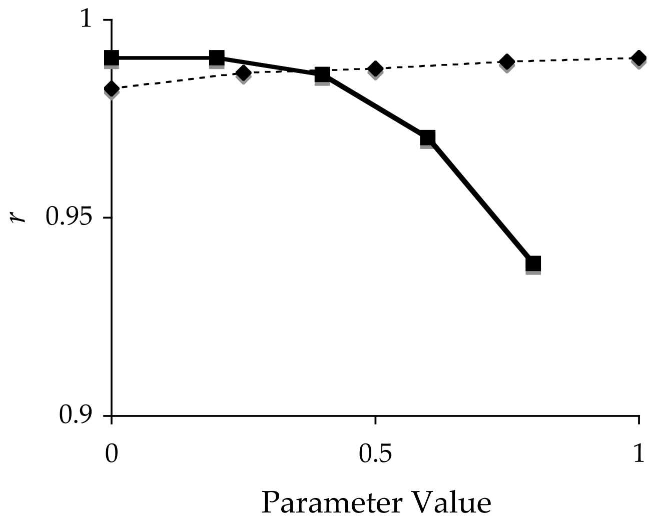

Figure 4.7: As the interaction parameter β (dotted line) is increased, a better fit is achieved with normalized human data. As the confidence threshold α (solid line) is increased, the opposite holds.

of atypical-competitive conditions. In a correlation of corresponding time steps, these normalized x-coordinate proportions show a strong relationship with **M** activation in the high interaction ($\beta = 1$) and early motor input condition ($\alpha = 0$; $r = .99$, $p < .0001$). When conducting the same regression analysis over the other parameter values, one obtains increases in fit in the expected directions (see Fig. 4.7). The best fit with human data is obtained when interaction between integration and motor layers is maximized: When β is large (high interaction between **I** and **M**), and α is small (early input to **M**).

Discussion

Results of this simulation suggest that a model of categorization which involves interaction between decision-making and response-generation components fits best with data from human participants. In addition, when the model involves early activation of the response component by the decision component, the pattern of competition seen in the simulation conditions again matches human data. Put simply, the data observed in human experiments are modeled by a system in which decision and action fluidly interact.

It is important to note a few caveats before generating a prediction from the model. First, as noted in the introduction to the normalized-recurrence architecture, this simulation is not intended to capture categorization in full detail. Instead, the architecture instantiates parallel feature processing that, through iteration, leads to a stable network state. The specific configurations employed in the current simulation resemble a simplified semantic feature space for animal categories and their exemplars. Second, the parameter range explored is not intended to map directly onto any neurophysiological variables. The fact that the best match the human data are the parameters' extreme values by itself makes it inappropriate to pursue this -- in fact, given the simplicity of the model, seeking such a correspondence may be setting it too great a task. A related concern is the exact nature of the neural processes governing manual output. It may be possible that multiple independent motor programs generate the response dynamics in such experiments as Spivey et al. (2005) and Dale et al. (in press). The model cannot currently address this issue.

Therefore, rather than capturing categorization and its underlying neural substrate, the model provides a simplified computational system that explores the effects of two interacting systems. One system acts through constraints imposed by features to decide on a relevant category node; another acts through this decision component at a given later time to mediate between only two possible stable states (the correct vs. incorrect node). Results suggest that, in such a context, fluid interaction between these components generates dynamic output patterns resembling human data.

Simulation 2: Typicality Gradients in Output

If this architecture is to serve as a basis for comparing the role of decision and action components, then it should generate predictions in different stimulus contexts. One explored here is a natural consequence of the feature-based processing of the network. As in numerous experiments in categorization (discussed in the introduction), the model should exhibit patterns of responses whose characteristics lie along typicality gradients: Response dynamics for *cat* should reveal less competition than that for *whale*, with graded results between these typical and atypical extremes.

Procedure

We selected mammals as the category whose typicality gradient will be explored in this simulation. Exemplars of this category were generated by the following procedure. Starting from a prototypical feature arrangement, we selected one or two feature layers, and randomly changed their value to correspond to that expected by another category. For example, we might change the prototypical mammal pattern by choosing LIMBS, and changing the initial network condition to begin with <*fins*> rather than <*legs*>. This generates approximately 50 atypical (sometimes hypothetical) exemplars. The response options activated with .5 at the beginning of a simulation run were the correct category (*mammal*), and a randomly selected incorrect category.

While the previous simulation used three groups of network condition, the current simulation involves a gradient of typicality. Rather than grouping the various generated exemplars, we seek to establish a relationship between typicality values (semantic distance; see below), and dependent measures drawn from a network's behavior when processing these exemplars. We investigated a network's response dynamics using two measures drawn from the behavior of **M** over multiple iterations of a run. The first is the total amount of time required for the motor units to achieve a maximum value of .95. This output measure corresponds to the amount of time the motor response is "in motion" before completion, as in human experiments in Spivey et al. (2005) and Dale et al. (in press). The second measure is the amount of activation achieved by the fifth time step. This value was chosen to sample, on average, network progress at approximately 40-50% of its total response. This serves as a crude measure of the "velocity" of the output dynamics, used similarly in Spivey et al. (2005) in human participants. Interaction and threshold components were at their maximum value ($\beta = 1$; $\alpha = 0$).

The response dynamics of the network is sensitive to typicality gradients if we observe a significant relationship between the similarity (or dissimilarity) of the modified exemplar to the prototypical mammalian feature pattern, and our output measures. A "dissimilarity" or distance measure was obtained using the Euclidean distance in feature space: Once the atypical vector was generated by one or two modifications, the Euclidean distance between it and the original prototypical mammalian vector was computed. After all network runs were completed, we conducted regression analyses between output measures, and exemplar distance from prototype.

Results

Regression results show a significant relationship between distance in feature space and both output measures. Total response time was positively related to typicality: The more typical an animal (shorter semantic distance), the less time required for the motor response to be generated ($r = .30$, $p < .001$). In addition, the activation of the network by the fifth time step was significantly related to distance from the prototype. Animal exemplars at a greater distance from the prototype had lower activation ($r = -.29$, $p < .001$).

We ran this same typicality-gradient simulation over parameter values that ranged from highly interactive between decision and action. The highest, used in the first regression analysis, has the **M** to **I** parameter (β) at a high value (1), and decision threshold (α), at a low level. We varied these values by jointly decreasing β and increasing α, and with each variation conducting regression analyses to test for typicality gradient effects. Results are presented in Fig. 4.8. As interactivity between decision and action components diminishes, typicality gradient effects substantially diminish.

Discussion

If response dynamics flows systematically from their cognitive processes, the current simulation predicts that typicality gradients should be exhibited in

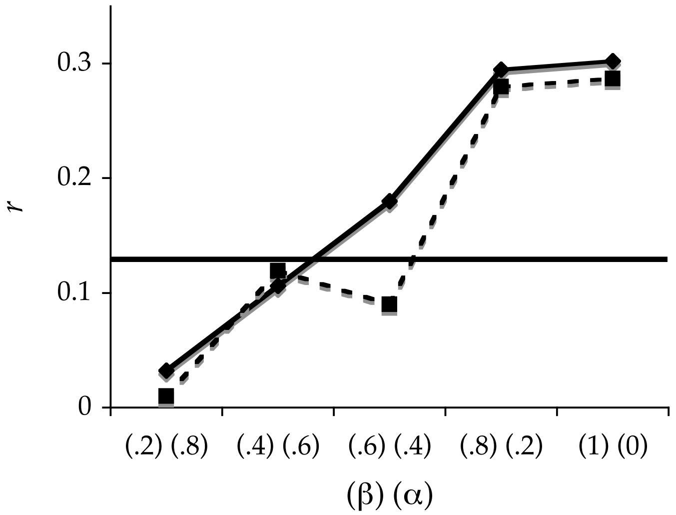

Figure 4.8: Time required for response (solid line), and motor activation of correct option (dotted line) correlated with Euclidean distance of generated exemplar as interaction (β) and decision threshold (α) vary. Significant typicality gradients are achieved as interaction grows (high interaction, low threshold for response). Note: *r* in the motor activation is a negative relationship -- i.e., higher the motor activation, the lower the distance to mammalian prototype. Solid horizontal line indicates cutoff for significant *r* given N.

measures drawn from output. These gradients do not hold when the model has little interaction between decision and action components. Like research on lexical decision and frequency (e.g., Abrams & Balota, 1990), this suggests that similar correlations between output and typicality measures will obtain in human experiments on categorization.

General Discussion

Normalized-recurrence provides a simple yet ideal arena for exploring the interaction among constraints that underlie taxonomic categorization. These constraints involve the parallel processing of features, and an accumulation of information that guides the categorization decision and resultant response. This model indeed resembles several recent simulations that also seek to characterize the time course of categorization (e.g., Lamberts, 2000). While the current model is simple, it offers some explicit insight into the possible relationship between the categorization decision, and the resultant response itself. In the model, patterns of competition observed in human data are obtained when both the decision process and the motor output concurrently and continuously interact. The kind of approach used here may suggest ways in which existing models (e.g., Nosofsky & Palmeri, 1997; Lamberts, 1995, 2000) might integrate action parameters, thereby providing a fuller picture of the cognitive process: Accounting for not just the hypothesized internal processes, but information accumulation all the way into the observable response behavior and its time course. For example, the Nosofsky and Palmeri model involves a random-walk process where exemplars race each other to help categorize a test exemplar. Including a motor component may simply involve integrating a second random-walk process in which possible category responses race each other for selection. Interactive parameters could then be similarly explored.

One possible promissory note about such models is a synthesis of "pre-decision" process models, and the experimentally observed relationship between process and response characteristics. Growing research on the embodiment of cognition (e.g., Barsalou, 1999; Clark, 1977) shows that parameters of action can impinge on the cognitive processing eventually leading into it (e.g., Glenberg & Kaschak, 2002; Tucker & Ellis, 1998). The current model provides some simple means by which computational mechanisms of these effects can be devised.

The current model has some obvious limitations. Firstly, it is extremely simple. The model simply embodies basic computational principles regarding constraint-based feature processing, and the continuous integration of this information. However, it is perhaps surprising that such simple assumption can succeed in generating the kind of response patterns observed in the human experiments: Both competitive and noncompetitive trials and their graded response characteristics can be captured in this system.

Secondly, as discussed earlier, the model does not seek to resolve current debate regarding opposing theories of the categorization process. While the model may prima facie have difficulty fitting knowledge-based results of categorization, its properties serve as an exploration of the time course of categorization given certain conditions. Those assumed here at least sufficient to model basic categorization tasks. While this is promising, others seek to scale up computational models of categorization and category learning (Love et al., 2004). The current model may again provide some motivation for integrating output-based dynamics to model responses.

In conclusion, these simulations aim to further the argument that the process of categorization does not serially project into the effectors. Rather, manual output responses may receive continuous information flow from the categorization process as it unfolds. Moreover, these action parameters themselves may serve as further constraints on the categorization process --

thereby suggesting that cognition and action are not bounded by strict demarcation, but interact fluidly and systematically.

CHAPTER FIVE

Typicality Gradients in Lexical Categorization Revealed by Graded Manual Responses

Introduction

An increasing amount of research reveals that dynamic characteristics of motor output reflect underlying cognitive processing, rather than simply reflecting the discrete decision resulting from that processing. For example, when the cognitive system directs manual output amidst an array of graspable objects, the arm's movement does not always proceed in ballistic fashion toward a single selected object, but may reveal subtle dynamic characteristics depending on the nature of underlying processing. Both manual output and oculomotor responses demonstrate these dynamic characteristics intrinsic to the temporal extent of a response, not just the final outcome of the response. For example, Doyle and Walker (2001) demonstrate that saccadic eye movements reflect attentional processing of visual cues in a simple fixation experiment. Saccade trajectories to the same location exhibit very subtle differential curvature depending on the position of distractor or cue stimuli (see also Sheliga, Riggio, & Rizzolatti, 1995). Additionally, considerable research over the past 10 years has shown that eye movements offer a semi-continuous measure of ongoing cognitive processing (Ballard, Hayhoe, & Pelz, 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Underwood, 2005). Aggregate data from eye movements often indicate a graded nature inherent to cognition in general.

Similar findings demonstrate that *manual* motor output can reveal graded representations. The force and velocity of manual responses vary concomitantly with frequency in a lexical decision task (Abrams & Balota, 1991; Balota & Abrams, 1995), and response and stimulus probability in simple reaction-time tasks (Mattes, Ulrich, & Miller, 2002; Ulrich, Mattes, & Miller, 1999; see also Osman, Kornblum, & Meyer, 1986; Balota, Boland, & Shields, 1989). And in experimental work similar to the saccade trajectory experiments described above, Tipper, Howard, and Jackson (1997) have shown that arm trajectories can curve depending on the visual distractor context in which reaching motions are made (see also Tipper et al., 1992; Sheliga et al., 1997). More recently, Spivey, Grosjean, and Knoblich (2005) and Dale, Kehoe, and Spivey (in press) used computer-mouse trajectories to show that graded manual output reveals temporal continuity in the underlying cognitive processes in spoken word recognition and categorization.

In the latter two studies, manual trajectories were measured through streaming x-y coordinates of computer mouse movement, and revealed attraction to other response choices in the visual display. For example, in Dale et al. (in press), mouse trajectories were recorded during lexical and perceptual categorization of animal exemplars. Participants categorized an animal by clicking the mouse on one of two category choices. Mouse-movement trajectories consisted of a movement from the bottom center of the screen, to the correct target on the upper left- or right-hand corner of the screen (beside which was a competing category label). Target trials used atypical animals (e.g., *whale*) with an incorrect competitor category that had considerable overlap in terms of semantic and visual features (e.g., *fish*). Though participants responded by clicking the appropriate category (e.g., *mammal*), mouse-movement trajectories exhibited substantial attraction toward the competitor category. Competing activation of the incorrect category in these trials was evident even in the properties of the resultant motor output, and not simply in the decision processes leading up to it. Information flows from the sensors into the categorization process and does not "discretize" before issuing motor output to the effectors. Instead, the effectors themselves seem to reflect some of this processing given the typicality of the exemplar (e.g., *rabbit* vs. *whale*), and the featural overlap with the competing category (e.g., *fish*).

This literature most often explores movement dynamics in simple experimental manipulations. For example, in both the saccade (e.g., Doyle & Walker, 2001) and manual response research (e.g., Tipper et al., 1997), motor trajectory curvature occurs in simple contexts containing visual distractors. An important outstanding concern is the extent to which properties of motor dynamics reflect finer-grained aspects of the underlying processing task. To make this point clearer, consider the categorization experiments in Dale et al. (in press) just described. While these experiments relied on two groups of animal exemplars, highly typical (e.g., *rabbit*) and highly atypical (e.g., *whale*), they did not explore whether typicality gradients between these extremes are evident in mouse trajectories. Similar research on lexical decision suggests that there should be a relationship between such stimulus parameters and motor output (e.g., Abrams & Balota, 1991).

Given the extensive influence of typicality in categorization (see Murphy, 2002), and that motor output may reflect cognitive processing, one should also expect that motor output would reveal gradedness as a function of typicality in a similar categorization task. Even further details regarding the underlying factors contributing to graded motor output can be acquired by exploring what specific semantic features define the typicality gradients along which output varies. For example, when categorizing animal exemplar names (lexical items), one might expect that certain semantic features constraining that process would exert more of an influence than others. For example, perhaps static visual features, such as shape or color, exert more of an influence if animal exemplars were presented in picture form.

The following experiment aims to supply some insight into these issues. A large set of animal *lexical items* is categorized in the same task as Dale et al. (in press), but the competing, incorrect category is randomly selected from 4 possible alternatives. Motor output is again measured in terms of mouse trajectories. The subsequent analysis provides clues about finer-grained processing exhibited by motor output. Firstly, typicality-gradient effects should be revealed in the motor output. Secondly, because lexical items are being processed, we use multiple typicality measures to reveal that specific feature sets are related to the typicality gradient along which motor output varies. Results demonstrate that the effectors exhibit cognitive processing in systematic ways: Effects found in research on categorization decisions are also revealed in their motor output. In addition, motor output can uncover the semantic features of the stimuli that underlie the lexical categorization task.

Experiment

Participants

31 Cornell University undergraduates participated in the study for extra credit in psychology courses. All participants were right-handed.

Materials

Basic-level animal names were selected from the concept-name set of McRae et al.'s (1997) study in which participants listed features of various animals and objects. For the present study, we used 125 of McRae et al.'s animal names. Each animal corresponded with a superordinate category of mammal, fish, reptile, bird, or insect. The experiment was programmed using RealBasic, and presented on an Apple eMac computer. A standard one-button Apple mouse was sampled using RealBasic's Timer control at a rate of approximately 40 Hz.

Procedure

At the start of each experimental trial, participants were presented with two superordinate animal categories, one category name in the upper right-hand corner of the computer screen and one category name in the upper left-hand corner (with approximately 16 degrees of visual angle between categories). After 2000 ms, a 1 cm^2 square appeared at the bottom center of the screen (approximately 13 degrees of visual angle from either category name). When participants clicked on this square with the computer mouse, the square was replaced by a basic-level animal name that corresponded with one of the two superordinate animal categories already at the top of the screen. Participants' task was to click on the superordinate category corresponding with the animal name for that trial. Before the 125 experimental trials, each participant completed three practice trials. The animal name presentation order and each trial's *incorrect* category were randomized. Likewise, the presentation side of the category names (left vs. right) was also random. Streaming x-y coordinates were recorded between participants' click on the square, and their final categorization choice (see Fig. 5.1A).

Typicality/Similarity Measures

Three different typicality measures were explored. These measures were aimed at quantifying the proximity of each animal to the incorrect categories in semantic feature space using a distance metric. Typicality here refers to the proximity in semantic space of animals to categories of which they are *not* members (e.g., in semantic feature space, a whale is more typical of the category *fish* than of the category *cat*). Typicality here is then *(dis)similarity*, expressed through semantic distance, between the animal exemplar and the central featural properties of the incorrect, competing label which is assumed to draw the manual trajectories towards it.

Each of the 125 animal names can be represented as a sparse semantic vector in a 205-dimension feature space drawn from McRae et al.'s (1997) concept-name set. These semantic features were organized into three groupings: visual-dynamic, visual-static, and category features. Visual-dynamic features (109 total) depicted specific animal behaviors, e.g., "swims" and "eats seeds." Visual-static features (74) depicted specific appearance characteristics, e.g., "has a long tail" and "is furry." Category terms (22) included non-visual labels often used to classify animals, e.g., "is domestic" and "is endangered." These groupings defined three typicality measures: Proximity in semantic space using dynamic features, static features, and the full 205 feature set.

The 125 animals can be mapped in a semantic space with dimensionality of the number of relevant features (dynamic, static, or full).

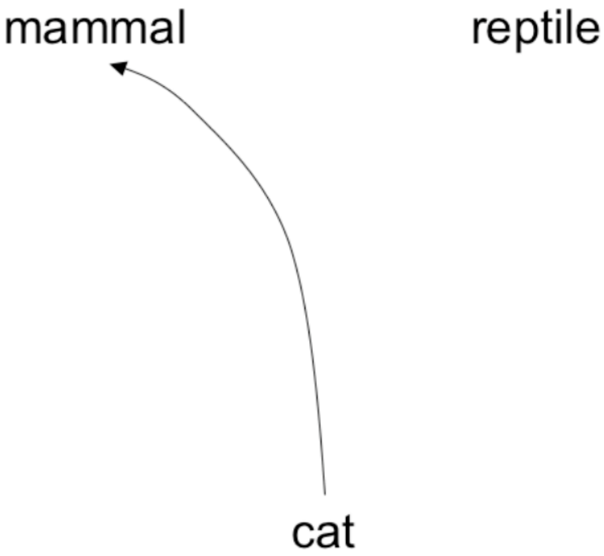

Figure 5.1A: What participants saw, along with a hypothetical mouse trajectory to the correct category.

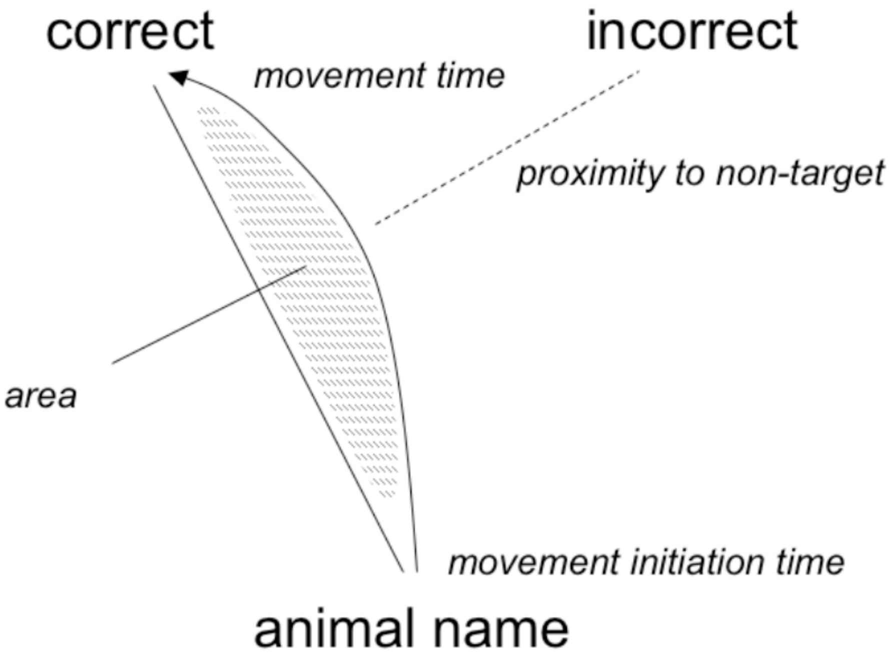

Figure 5.1B: Depiction of dependent measures.

Each category's central tendency point was determined by averaging the coordinates of its constituent animals. In the resulting space, the most typical animal exemplars of each category clustered around their category's central tendency point. On the other hand, atypical animal exemplars (e.g., *whale*) of each category were positioned much further away, often nearer to the central tendency point of an entirely different animal category (e.g., *fish*).

Output Measures and Analyses

Streaming x-y coordinates were sampled from the presentation of the lexical item, to the final categorization click, and only correct trials were subjected to analysis. Four properties of manual output were extracted from these mouse-movement data (see Fig. 5.1B). First, the movement initiation time was computed by measuring the number of samples before mouse movement was detected (i.e., while the cursor was motionless after the start of a trial). From the remaining trajectory representing output motion, we calculated the number of time steps required to finish the categorization (movement time), the total area occupied by the trajectory compared to an assumed straight line to the correct category (area), and the closest point in the trajectory to the competing category (proximity to incorrect category). All measures were based on raw Timer control samples extracted in RealBasic, rather than an approximated temporal measure.

Analyses aimed to detect the extent to which each feature set predicted these dependent measures. In other words, if an animal's feature vector is close to a competing category's vector, one would predict that output measures would represent attraction towards that category label. Therefore, proximity to the incorrect category label should be smaller (closer to label) if the distance between the animal and incorrect category in semantic space is small. To test this, we performed a two-part analysis. Both tests make use of item-based observations. Each animal is paired with 4 randomly selected non-targets. With 125 animals, we therefore have 500 item types supplied by the experiment. The subsequent analyses are based on these 500 item-category pairs. Each pair has observations averaged across participants who encountered it in a trial during the experiment. These data were used in the two-part analysis.

In the first part, we separated animals along typicality gradients by using one standard deviation (SD) of the mean distance from animals to a given category (i.e., all animals to *bird*). This produced two groups of animals for each category. The first group (N ≅ 60 in the three feature sets), one SD *below* the mean distance, represents those animals *close in semantic space* to the competing category. The second group (N ≅ 60), one SD *above*, is particularly distant from the central tendency of the category (see Table 5.1 for examples). We conducted *t*-tests between these groups for each measure.

In a second test, we sought to confirm that significant relationships between typicality and the measures held across the entire group of animals. To do this, we computed a regression coefficient between the typicality measures and the output measures. Proximity to incorrect category, for example, should reveal a positive relationship with typicality (measured here in distance): Closer semantic distance measures should significantly predict closer spatial proximity measures in the mouse movements.

Table 5.1: Some examples of close and distant animals in *dynamic* semantic space

Close animals (one SD below mean of distance for given category)

animal	category	distance
seal	fish	.30
dove	insect	.31
walrus	fish	.34

Distant animals (one SD above mean of distance for given category)

animal	category	distance
skunk	insect	1.09
moth	fish	1.08
bull	bird	1.06

Results

Participants erred on 2.45% of experimental trials. These trials were not included in statistical tests.

In comparing the groups of above/below one SD to the mean distance in *dynamic* feature space, animals close to competing categories exhibited larger trajectory area (69386.7 vs. 64671.4 pixels^2, $t(122) = 2.0$, $p < .05$), longer time in motion (26.3 vs. 23.3 samples, $t(122) = 2.6$, $p < .05$), and significantly closer proximity to the competing category label (340.7 vs. 368.1 pixels, $t(122) = 2.6$, $p < .05$). Neither visual nor full 205-dimensional feature space exhibited any significant or marginally significant differences.

Regression analyses revealed the same pattern. Only dynamic feature space again revealed significant relationships between output measures and distance. These are presented in Table 5.2, along with the results for static and full space regressions.

Table 5.2: Regressions across sets and output measures

	r		
Measure	Dynamic	Static	Full
Area	-.13**	-.01	-.03
In motion	-.15***	.04	.00
Proximity	.13**	.04	.04

Discussion

Results further support that dynamic properties of motor output reflect cognitive processing. The output measures significantly relate to typicality gradients for categories and animals that are not their members: The closer the proximity in semantic space between animal and incorrect category, the greater the attraction of the manual response towards that category label. In addition, by separating the feature space in terms of different semantic content (function vs. visual features), we find that the dynamic feature set predicts attraction to the non-target label, while static visual features do not. Although there are a number of caveats regarding the immediate implications of these semantic spaces (see below), this at least suggests that motor output is reflecting finer-grained featural semantics that underlie lexical categorization in the task.

An additional analysis that may test this claim is to conduct similar tests using Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), a semantic representation scheme based on co-occurrence of lexical items in text. If motor output reflects the processing of lexical items specifically, then we should also observe significant results when computing typicality gradients in terms of LSA measures.

LSA Analysis

LSA measures computed semantic *similarity* (rather than distance or dissimilarity) between animal names and the 4 alternative categories to which they do not belong. We should therefore expect the reverse pattern of results for our dependent measures.

Exactly the same strategy was used to separate two groups of differing distance from mean typicality. These did not produce significant results. However, unlike the regression results above for visual, but similarly for the function feature space, LSA significantly predicted all dependent measures: area, $r = .14$, $p < .01$, movement time, $r = .11$, $p < .05$, and proximity, $r = -.10$, $p < .05$. While the SD separation of animal-category pairs did not attain significance, the regression results reveal that typicality gradient effects hold with LSA measures.

Movement Initiation Time

We present a final analysis that tests a prediction made by the perspective that processing flows into the effectors. If cognition indeed does not discretize information prior to initiating motor output during categorization, then there is likely an important temporal component to the process. If a participant allows a relatively large amount of time to pass during a trial before initiating their response, then one would not expect there to be significant dynamical competition in output: By spending more time evaluating the animal name and category labels before moving, the decision process may reach a higher level of certainty. The upshot may be a more reliably linear, ballistic movement to the correct category label.

We looked at the relationship between movement initiation time and the output measures. One measure exhibits a significant relationship. Proximity to competitor is positively related ($r = .15$, $p < .001$). In other words, the longer the amount of time spent before initiating motor movement, the less spatial attraction exerted by the competing category.

General Discussion

In everyday life, our arms move continuously during such tasks as gesturing in conversation, organizing objects on a table, or managing cooking ingredients. Their neural substrate is a fairly slow system (relying heavily on prediction; e.g., Flanagan & Lolley, 2001), not firing off movements in staccato fashion (much like saccades), but often changing course mid-path, or issuing graded movements as it directs the arms to their target. This intuition about everyday movement is demonstrated in the foregoing results, and in the array of motor-dynamics findings reviewed above. Even in a relatively "higher-order" cognitive process such as categorization, manual output has internal characteristics that likely reflect the categorization process itself. In the above results, mouse trajectories vary concomitantly with typicality gradients, and these gradients may lie along semantic dimensions relevant to the processing task (*lexical* categorization).

Nevertheless, a number of important limitations should be noted. First, the results, while robust, are thus far fairly weak. There may be a number of reasons for this. Firstly, Dale et al. discovered that lexical stimuli generate weaker competition effects in mouse trajectories -- if picture stimuli were used, stronger effects of typicality may be observed, along with the potential importance of static visual features.

Secondly, previous findings with saccadic trajectories show an effect of location of distractors relative to targets, resulting in varying strengths of trajectory curvature (see Godijn & Theeuwes, 2002, for a review). In the kinds of experiments reported here, it is uncertain where or whether there are effects of relative location. Further studies may explore different locations of competing category labels, and whether this weak result is inherent to the nature of interaction between cognition and action, or perhaps the design presented here involved response choices situated too close or too far to reveal more marked trajectory effects.

Thirdly, little was done to transform the semantic feature space afforded by McRae et al.'s (1997) concept set (e.g., multidimensional scaling, row/column normalization, similarity-metric transformation). Also, these feature sets were not intended to define categories – so raw feature values in Euclidean space were used to infer category clusters. Moreover, we did not make use of typicality scores for correct categories. We feel that this is in fact a more conservative test of the predictions made above, because raw Euclidean distance between animal and incorrect category relates to motor measures, without adding the additional information regarding proximity to correct category semantics. Further detailed analysis of McRae et al.'s (1997) semantic feature space may thus strengthen these results.

Despite these limitations, the proximity of correct and incorrect category labels did produce dynamic motor movement effects that reliably correlate with raw semantic feature space. The results further contribute to a wide literature on processing distinctions between mode of stimulus presentation: Categorization of lexical items may rely on semantic information that is distinct from that centrally involved in categorizing pictures of animal exemplars (e.g., Snodgrass, 1984; Viswanathan & Childers, 2003). Although it is possible that the limitations may have rendered detection of visual static effects undetectable, further exploration may seek to explore the contribution of visual (or other perceptual) information in both decision- and output-based measures in lexical categorization (see, e.g., Pulvermüller, 1999). One approach is to make use of picture stimuli of the 125 animal names (e.g., Dale et al., in press). Another, as mentioned, is to subject the semantic feature space used here to more detailed analyses, perhaps revealing the relevance of other visual or other perceptual semantic features in more sensitive tests.

The findings reported here challenge the common intuition that the properties of motor output are uninformative of cognition. Perhaps more importantly, they suggest that processing flows in systematic ways into motor behaviors, rather than simply being collapsed onto them to generate a categorical response (cf. Gold & Shadlen, 2000). They may recommend a "cascadic flow" perspective on cognition that sees information flow continuously from sensors to effectors (McClelland, 1979; Balota & Abrams, 1995; Spivey et al., 2005). These results lend support to this perspective, which challenges the discrete perspective on the way in which a central cognitive process, such as categorization, becomes action.

CHAPTER SIX

From Apples and Oranges to Symbolic Dynamics: A Framework for Conciliating Notions of Cognitive Representation

1. Introduction

Since its inception, cognitive science has offered up a wide array of hypothetical constructs, intervening somewhere between our sensors and our effectors, to explain our observable behavior. Many of these constructs can be filed under the umbrella term "representation." Representations might "stand for" things in the world (Bloom & Markson, 1998), be asymmetrically dependent with worldly objects (Fodor, 1987), they might get stored or processed or recalled (Atkinson & Shiffrin, 1968), and they surely change somehow during development and learning (Danovitch & Keil, 2004). This generic construct has appeared and reappeared in a variety of forms, labeled variously with the terms "traces" (e.g., Rosen, 1975), "schemata" (e.g., Bartlett, 1932; Neisser, 1976), "categories" (e.g., Rosch, 1975), "concepts" (e.g., Medin, 1989), "object files" (e.g., Feigenson & Carey, 2003), and so on – perhaps describable as different *forms* of representation.

There is no single agreed upon theory or definition of representation among cognitive scientists (Dietrich & Markman, 2003). The details about any particular brand of representation are mostly specific to the theory in which it plays a role – but each brand can be characterized in terms of some basic features. Nonetheless, even these most fundamental properties of representation are the subject of continuing debate in cognitive science. One such property concerns the temporal and spatial extent of representational states. There are two basic sides to the traditional version of this debate. One family of theories may be described as "discrete-symbolic," because they claim that internal representational states involve discrete computational information structures that are manipulated in logical algorithmic processes. Here, "computational" can be understood intuitively as structures that take the form of something a digital computer would process – content that is discrete in space and time. A competing family of theories may be described as "continuous-distributed," because they instead invoke representational states that are spread out in space, and extended in time. These states are graded, statistical, and probabilistic – they cannot be individuated discretely in time, or uniquely in their informational content. Continuous-distributed representations contain probabilistic informational patterns that might blend into other such representations, whereas discrete-symbolic representations are by definition independent uniquely identifiable states that are each separate from, yet used in conjunction with, other discrete representations (Dietrich & Markman, 2003).

Debate continues about which composition is the most appropriate foundation for cognitive explanation. Recently, Dietrich and Markman (2003; Markman & Dietrich, 2000) have offered persuasive arguments about the crucial role of symbolic representation in higher-order cognition, such as conceptual organization, problem solving, and language (see also Marcus, 2001; Pinker, 1997). Meanwhile, Spivey and Dale (2004) argue that a continuous composition is extensively evidenced throughout even complex cognition, offering examples from real-time language processing and visual cognition (see also Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Port & Van Gelder, 1995). In this article, we provide a review and discussion of a mathematical terrain in which these two representational formats can be directly compared and evaluated. We suggest that a kind of "mathematization" of the problem space, in terms of nonlinear dynamical systems and symbolic dynamics, can aid in a variety of ways. The descriptive power of dynamical systems, and the computational power of symbolic dynamics based on them, can reveal an epistemological synthesis of this debate, and offer an illuminating framework for exploring such conceptual conciliation. For reasons that we describe later, we do not expect the framework of symbolic dynamics to make moot the debate between discrete and continuous descriptions of mental activity, but rather may pose as the level playing field on which the debate may actually achieve a consensual resolution.

A case may be made for mathematization of scientific domains as a course toward resolving theoretical disputes, clarifying conceptual confusions, and making potential decisions concerning the greater validity of one verbalized scientific description over another (e.g., for a discussion of this in psychology, see Meehl, 1998). What early calculus did for Newtonian mechanics, tensor calculus for general relativity, symbolic logic for computer science, among other possible examples, is to provide a formal framework for exploring the relationships among observables, thereby making explicit predictions that can be tested empirically. A common mathematical framework, within which different theories compete, permits more rigorous evaluation of hypotheses that otherwise would be couched merely in a verbalized form. For example, despite the growing popularity of quantitative models, it is not difficult to find theories in cognitive psychology whose existence enjoys only verbal description. This certainly does not invalidate the potential contribution of these theories – observers should always be reminded of the youth of cognitive science. The bigger problem is that multiple competing theories in verbal form may be conducive to debate with little chance of resolution. For one thing, without a common formal framework, it may be difficult to tell if two competing theoretical schemes are in fact mutually exclusive, or perhaps even extensionally equivalent. In other words, without explicit formulation of the relationship among theoretical entities, in more or less formal terms, it may be difficult to determine whether two competing entities are two distinct incommensurable accounts, two different aspects of one process, or merely two different descriptions of the same process. Secondly, theorists of differing persuasions may be talking past one another, preventing a Hegelian "thesis-antithesis-synthesis" resolution that may be revealed by a common framework permitting conciliation of competing notions.

There are pursuits in cognitive science that benefit from aspects of formalization at present. For example, connectionist models have been used as a common information-processing framework for evaluating competing theoretical accounts of cognitive processes involved in language. McRae, Spivey-Knowlton, & Tanenhaus (1998) used a localist attractor network to compare directly the immediate information-integration predictions from the constraint-based theory of sentence processing (MacDonald, Pearlmutter, & Seidenberg, 1994) to the architecturally delayed integration predictions from the Garden-Path theory (Frazier, 1995). Drawing from work by Elman (1990) and Schütze (1994), Spivey-Knowlton and Saffran (1995) used a connectionist-like framework to directly compare the advantages of incremental prediction and explicit negative evidence in learning a simple language structure. Also, particularly relevant to our concerns here, homogeneous versus hybrid simulations using connectionist principles have recently been developed to compare dual- vs. single-route models of reading (Harm & Seidenberg, 1999; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001). Other formal frameworks, such as Bayesian modeling (Tenenbaum & Griffiths, 2001), genetic algorithms (Chater, Christiansen & Reali, 2004), and statistical models of sentence processing (Chater & Redington, 1996) have been manipulated in ways that allow comparison of competing theories.

The overarching theoretical concern, however, is that many of these models involve too many degrees of freedom to make them a sufficiently agreed-upon common ground for comparing theoretical constructs whose properties are highly disparate. Comparing individual models of particular processes is surely valuable and inevitable, yet fundamental theoretical differences in cognitive science cannot be contacted through manipulating already-existing architectures.[1] The debate over representation is particularly illustrative in this respect. Those who propose symbolic rules and representations have often urged core qualitative differences between these

---

[1] This is not to say that theoretical debate cannot proceed by comparing distinct architectures and their ability to capture the data – because this is already an area of productive debate in cognitive science (e.g., Pitt, Myung, & Zhang, 2002; Roberts & Pashler, 2000). We are instead recommending the use of a single formal framework that permits comparison of different theoretical constructs that could exist within that framework. The relative contribution of these different constructs for fitting experimental data, within the same set of agreed mathematical or formal principles, would then adjudicate between the competing theories.

kinds of states and the probabilistic distributed states that are the hallmark of statistical models, such as connectionist simulations. For this reason, choosing a formalization that has a pre-existing affiliation with a particular theoretical framework, such as a production system or a connectionist model, biases the enterprise toward the theory from which the model originated. A common ground should instead derive from a formalization that can already adequately incorporate and implement both sides of the theoretical debate. A mathematization or formalization of the debate over representation needs a common framework for directly comparing symbolic states and dynamic processes within the same explanatory arena.

In this article, we propose that a branch of dynamical systems theory may serve as this common ground. Symbolic dynamics has both complemented understanding of the continuous-time nature of systems, along with providing groundbreaking insight into the computational power inherent in dynamical systems (e.g., Crutchfield, 1994). A few proposals concerning symbolic dynamics have already been offered from contributors outside of cognitive science. Below we introduce symbolic dynamics for the cognitive scientist, and review some of these proposals. Before describing this framework, we first offer discussion supporting the position that some portion of our perceptuałcognitive processes is already awash in continuity: That the best physical description of the mind/brain must invariably invoke, at some level, continuous (or discretely-approximated continuous) bases for understanding the substrate of cognition. A theory of cognition is superimposed on this continuity in two broad ways historically: Discrete-symbolic or continuous-distributed representational states and processes as the theoretical basis for cognitive explanation. We then introduce symbolic dynamics as a framework that can incorporate (and, in certain special cases, show the equivalence of) both kinds of explanation. We finally offer our own consideration of symbolic dynamics, with its potential contribution to and limitations in cognitive science.

2. Continuity

A key point Dietrich and Markman (2003) use to support discrete representations is our cognitive system's ability to form categories for objects in our world: "If a system categorizes environmental inputs then it has discrete representations" (p. 101). Moreover, they argue, continuous accounts of categorization would miss the mark, since categorization by definition involves consistent responses to completely distinct elements in our environment – it makes no room for continuity (historically this may be arguable, cf. Rips, Shoben & Smith, 1973; Rosch, 1975). Although Dietrich and Markman offer extensive discussion to forestall possible replies, there remains a problem with this perspective. What the authors dub "enduring classes of sameness" (p. 101), in an environment that our system must categorize, involve discrete internal representations whose primary evidence comes from what might be called *time-course irrelevant* responses during a cognitive task. An outcome-based response measure, such as a forced-choice categorization task, is time-course irrelevant because the temporal dynamics of representational activations leading up to the forced choice go undetected by the response measure. Such response measures may artificially exaggerate the degree to which the enduring classes exhibit their sameness. For example, even the cognitive literature's darling of discrete mental events, categorical speech perception (e.g., Harnad, 1987; Liberman, Harris, Hoffman, & Griffith, 1957), exhibits some graded sensitivity to continuous phonetic feature information when its temporal dynamics is measured with reaction times and eye movements (McMurray, Tanenhaus, Aslin, & Spivey, 2003; Pisoni & Tash, 1974; see section 2.2 below). In this article, we consider two realms, visual cognition and language comprehension, in which an unmistakable continuity is observed even in seemingly discrete, categorical tasks (see Spivey & Dale, 2004, for further review).

2.1 Vision

Vision research is replete with examples of continuity in real-time perception. The gradual settling of a population code of neurons, over the course of hundreds of milliseconds, is a typical way to think about how the visual system recognizes objects and faces. Compelling visualizations of the continuous manner in which sensory input gradually produces a percept can easily be found in visual neuroscience. We briefly consider three cases: Object and face recognition, visual search, and perceptual decisions.

Rolls and Tovee (1995) recorded from neurons in macaque inferotemporal cortex, and found that it takes a few hundred milliseconds for a responsive population of cells to achieve their appropriate firing rates indicating full identification of a fixated object or face. The cumulative information (in bits) provided by an inferotemporal neuron in the service of recognizing a face or object accrues continuously (though nonlinearly) over the course of about 350 milliseconds until asymptote. Perrett, Oram, and Ashbridge (1998) demonstrate similar patterns of gradual accumulation of neuronal evidence during face recognition. When an object or face is partly rotated away from the frontal view, recognition or matching will generally take longer as a function of how far it is rotated (e.g., Cooper & Shepard, 1973; Jolicoeur, 1985; Shepard & Metzler, 1971; see also Georgopoulos, Lurito, Petrides, Schwartz, & Massey, 1989). Perrett et al. (1998) describe recordings from cells in the monkey temporal cortex during viewing of frontal, three-quarter profile, profile, and quarter profile schematic faces. When the accumulated action potentials are plotted over time, these curves gradually rise to asymptote over the course of several hundred milliseconds, but at different rates as a function of how canonical the face orientation is.

The same kind of gradual accumulation of perceptual evidence can be observed when multiple objects are competing for attention during visual search. Although a serial-processing account has argued that the observer allocates attentional resources wholly and discretely to individual objects, one at a time (e.g., Treisman & Gelade, 1980; Treisman, 1988; see also Wolfe, 1992), a parallel-processing account is recently being developed in which attention is best characterized as involving partially-active representations of objects simultaneously competing for probabilistic mappings onto motor output (e.g., Desimone & Duncan, 1995; Reynolds & Desimone, 2001). In fact, a wide range of studies have been suggesting that the traditional distinction between putatively "serial" and "parallel" search functions is best revised as a continuum of search efficiency rather than two separate mechanisms of visual search (Duncan & Humphreys, 1989; Nakayama & Joseph, 1998; Olds, Cowan, Joliceur, 2000; Wolfe, 1998; see also Spivey, Tyler, Eberhard, & Tanenhaus, 2001).

Finally, Gold and Shadlen (2000) examined decision processes in macaque visual perception. A common task in visual psychophysics involves presenting a display of quasi-randomly moving dots. As the experimenter increases the proportion of dots that move in a roughly consistent direction, the perception of a coherent direction of flow amidst the dots becomes more apparent (Britten, Shadlen, Newsome, & Movshon 1992). Monkeys were trained to indicate the perceived direction of dot flow, upon offset of the stimulus, by making an eye movement to one peripheral location or an opposite one. After identifying a relevant frontal-eye field (FEF) region, electrical microstimulation evoked an involuntary saccade that was perpendicular to the two voluntary response saccades. On some of the direction-of-flow judgment trials, this region was microstimulated immediately after the moving dot display disappeared, i.e., exactly when the monkey was supposed to produce a voluntary eye movement that would indicate his response regarding the perceived direction of flow of the dots. By incrementally increasing viewing time of the stimulus before this microstimulation, the experimenters were able to observe the gradual increase in "strength" or "confidence" of the perceptual decision over time, as indicated by the degree to which that voluntary saccade "leaked into" the execution of FEF-microstimulated evoked saccade. Thus, the population of FEF cells that produced the evoked saccade were already somewhere in the process of settling toward a pattern of activation that would produce the voluntary response saccade. If the microstimulation took place early on in this decision process, rather little effect of the voluntary response would be apparent in the direction of the evoked saccade, but if the miscrostimulation took place later on in the decision process, a significant amount of the voluntary response would be apparent in the direction of the evoked saccade. This finding suggests that decision processes themselves may be coextensive with the gradual settling of partially-active and competing neural representations in oculomotor areas of cortex (Gold & Shadlen, 2001; Schall, 2000; see also Georgopoulos, 1995).

2.2 Language

Much like visual perception, language comprehension also exhibits a form of temporal dynamics that reveals underlying continuous-distributed formats of representation. There is considerable evidence that the multiple levels of linguistic complexity – comprehension of speech sounds, words, and sentences – are driven by graded, partially active representations. At the level of speech sounds, the phenomenon of categorical perception was long adduced as evidence for discrete representational states (Liberman, 1982). Lately it has been subject to extensive empirical investigation, and made consistent with more temporally dynamic approaches to categorization (Damper & Harnad, 2000; see also Anderson, Silverstein, Ritz, & Jones, 1977). For example, McMurray and Spivey (1999) tracked participants' eye movements while they performed the standard categorical identification task. This task involves categorizing sounds that lie on a voice-onset-time continuum between "bah" and "pah," by clicking a relevant icon on one or the other side of the computer screen. Thus, in addition to recording the participants' explicit choice, there was also a semi-continuous record of how the eyes tended toward one or the other response icon during categorization. With "pah" or "bah" sounds near the categorical boundary, eye movements exhibited conspicuous vacillation between categories before the overt mouse-click response was made. Despite the apparent categorical nature of the eventual choice, eye movements revealed a more continuous decision process that is sensitive to some of the graded acoustic-phonetic variation in the stimulus. These temporary phonemic ambiguities exhibit their effects not just in phoneme categorization tasks but also in spoken word recognition tasks (McMurray, Tanenhaus, & Aslin, 2002; McMurray et al., 2003).

At the level of word recognition, Spivey-Knowlton, Sedivy, Eberhard, & Tanenhaus (1994) demonstrated cohort effects in eye-movement patterns by having subjects follow instructions to manipulate real objects on a table. Participants sat in front of a table containing a central fixation cross and various objects around it (e.g., a fork, a mug, a candle). In some trials, objects whose names had similar initial phonemes were present on the table, available for manipulation (e.g., a bag of candy and a candle). Even before the spoken word was completed, eye-movements to both objects were often observed, such as briefly fixating the candle when instructed to "Pick up the candy." This phonologically similar object conspicuously attracting eye movements is indicative of the competing lexical representation being partially active during, and perhaps shortly after, delivery of the spoken word. Headband-mounted eye-tracking studies like this have demonstrated this real-time lexical competition using computer-displayed objects (Allopenna, Magnuson, & Tanenhaus, 1998), using artificial lexicons (Magnuson, Tanenhaus, Aslin, & Dahan, 2003), with young children (Fernald, Swingley, & Pinto, 2001), and even across two languages in bilingual participants (Marian & Spivey, 2003; Spivey & Marian, 1999).

Finally, in sentence processing, eye movements can again reveal the continuous intake and use of information during comprehension of a spoken utterance. For example, when presented with a real 3-D display containing an apple on a towel, another towel, and an empty box, and then instructed to "Put the apple on the towel in the box," participants often look briefly at the irrelevant lone towel near the end of the spoken instruction before returning their gaze to the apple, grasping it, and then placing it inside the box (Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In this case, the syntax is ambiguous as to whether the prepositional phrase "on the towel" is attached to the verb "put" (as a movement destination) or to the noun "apple" (as a modifier). Given the actions afforded by the display, the latter syntactic structure is the correct one. However, the brief fixation of the irrelevant lone towel indicates a temporary partially-activated incorrect parse of the sentence. To demonstrate the influence of visual context on this syntactic ambiguity resolution process, the display was slightly altered to include a second apple (resting on a napkin). In this case, the visual co-presence (in Herb Clark's, 1992, terms) of the two potential referents for the phrase "the apple" should encourage the listener to interpret the ambiguous prepositional phrase "on the towel" as a modifier (in order to determine which apple is being referred to) rather than as a movement destination (cf. Altmann & Steedman, 1988; Crain & Steedman, 1985; Spivey & Tanenhaus, 1998). And, indeed, with this two-referent display, participants rarely fixated the irrelevant lone towel, indicating that visual context had exerted an immediate influence on the incremental syntactic parsing of the spoken sentence (Spivey et al., 2002; Tanenhaus et al., 1995; see also Knoeferle, Crocker, Scheepers, & Pickering, 2003).

The current state of affairs in the field of sentence processing is at a consensus with regard to the continuity of information flow, and has been gradually approaching consensus with regard to the rapid integration of syntax, semantics, and pragmatic context (Trueswell & Tanenhaus, 2004). Just as the processing of speech sounds, at the scale of tens of milliseconds, appears to be characterized by multiple partially active phonemic representations competing over time (McMurray et al., 2002, 2003), and the comprehension of spoken words, at the scale of hundreds of milliseconds, appears to be characterized by multiple partially active lexical representations competing over time (Allopenna et al., 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986), so does the resolution of syntactic ambiguity, at the scale of seconds, appear to be characterized by multiple partially active syntactic representations competing over time (MacDonald et al., 1994; Spivey & Tanenhaus, 1998; Stevenson, 1994; Tabor & Tanenhaus, 1999).

2.3 Summary

From perception, such as visual processing, to cognition, such as the various levels of linguistic processing, there seems to be extensive evidence for continuous-distributed representation (see Spivey & Dale, 2004, for further discussion and examples). There nevertheless remains considerable debate about the nature of representation in other areas of cognition. In particular, in "high-level" cognitive processes such as reasoning and problem solving, there seems to be markedly slower success with continuous-distributed frameworks. This situation is exacerbated further by the comparatively rapid rise, and longer history, of discrete-symbolic accounts of reasoning and problem solving (e.g., Weizenbaum, 1966; Winograd, 1970; Newell & Simon, 1976).

If it can be granted that perception is largely driven by continuous change in processing states, then for the discrete-symbolic perspective to be right about cognition, there must be a “discretization” that happens somewhere in between perception and motor output. The debate can then be placed in the following terms: *How early in the system do our theories need to postulate this discretization, thus invoking a language of discrete symbols generated through causal influences of continuous processes?* A purely continuous-distributed account of cognition might place this discretization at the extreme end, only in between the motor action itself and its effects on the problem-solving environment. For example, although you may be trying to decide between moving your rook four squares up or three squares up in a game of chess, and this vacillation may even be visible in the continuous motor movement, in the end, only one of those alternatives actually happens. In contrast, the discrete-symbolic account of cognition urges an earlier discretization, recommending theories to work with symbolic states and algorithmic state-transition rules not long after perceptual processing. In such a case, the decision to move one’s rook three squares or four squares would be discretely made in an internal cognitive stage, and any vacillation observed in the motor movement would be best interpreted as a vestigial or epiphenomenal echo of the earlier temporary cognitive uncertainty.

If this formulation of the question is agreeable to both sides of the debate, then there exists a “common format of explication” that future research in high-level cognition might fruitfully use in order to consensually adjudicate between theories that propose an *internal* discretization of the brain's continuous dynamics and theories that propose only an *external* discretization of them. The mathematical arena of symbolic dynamics (e.g., Crutchfield, 1994; Devaney, 2003; Goertzel, 1998; Shalizi & Albers, submitted; see also Cleeremans, Servan-Schreiber & McClelland, 1989; Tabor, 2002, for related discussions) has exactly the ingredients for building systems that implement continuous temporal dynamics in a high-dimensional state space (of perception and of action) and can convert that continuous trajectory into an emitted string of formal logical symbols for describing external action-effects in a problem-solving environment, and also for describing internal cognitive states.[2] We next offer a very simple introduction to symbolic dynamics, and then discuss a number of issues relevant to its application in cognitive science.

3. Symbolic Dynamics

A continuous-distributed perspective on representation in perceptual and cognitive processes is often couched in model systems that change in time (be it continuous-time or discrete-time): dynamical systems. A dynamical systems framework provides a rich set of conceptual tools for cognitive science. The geometric entities in the study of dynamical systems can serve as an intuitive, and potentially mathematically rigorous, vocabulary for visualizing state changes within and between perception and cognition. As already mentioned, this strategy is widely used in many areas of cognitive science, and is often

---

[2] In fact, while defending discrete representations in cognition, Dietrich & Markman (2003) essentially describe the basic concept of symbolic dynamics, without referring to it by name, in their fourth argument (their discussion of Fig. 6.3).

considered its own framework for the study of cognition (Port & Van Gelder, 1995; Thelen & Smith, 1994; Ward, 2002; Kelso, 1995). In order to lay out this descriptive vocabulary, we briefly consider a simple iterative dynamical system, surely familiar to many readers, that illustrates a number of these geometric metaphors. Consider a function F(x) that maps real numbers onto real numbers by iteration: $F^2(x)$ is given by F(F(x)), and $F^3(x)$ by F(F(F(x))), etc. The logistic map is given by the equation

$$F(x) = \mu x (1 - x).$$

The time dimension is here represented by progressive iteration of the real value x into the function F, scaled by μ. The iterative process in this simple equation illustrates both stability, meta-stability, and transition into chaotic behavior. For example, when μ is between 0 and about 3.5, iteration of F(x) from any starting point of x will settle into stable attractor states – namely, the value of $F^n(x)$, as n becomes very large, stabilizes on one or more particular precisely-repeated values. These values are termed attractors in the logistic map's dynamics. As μ approaches about 3.6 or so, the logistic map exhibits chaotic behavior, where there is no stable attractor state, and its series of values can superficially appear random. One way of representing the transitions in state space of this system is through a phase plot, shown in Fig. 6.1. By tracking the value of x at each iteration, we can visualize the trajectory of the system from some random initial x into its attractor states (Fig. 6.2).

The logistic map is used extensively in textbooks on dynamical systems. Its curiosity lies in the rich complexity that emerges from iterating such a simple equation. In fact, the same issues considered for discerning the nature of the logistic map are a concern for innumerable systems of practical and theoretical interest. A wide variety of pure and applied mathematical techniques can be used to study these systems. Nevertheless, it is not uncommon that these methods can be outstripped by a system's complexity. A specific technique available to overcome such limitations, around for decades and gaining much attention of late, is termed symbolic dynamics, and offers a means of simplifying analysis (e.g., early on, Morse & Hedlund, 1938; see Devaney, 2003 and Williams, 2002 for review). A system's dynamics can be rendered symbolic by carving partitions or regions into its phase space, and assigning a unique numeric state or label to that partition. As the dynamical system's state changes in time, this trajectory is transformed into a sequence of emitted symbols corresponding to partitions in the space. Take, for example, the logistic map. We can represent its phase space as in Fig. 6.3, and divide the plot into two intervals, $I_0 = [0, ½]$ and $I_1 = (½, 1]$. When iterations of the system enter the first interval, the symbol "0" is output, and likewise "1" for the second. The dynamics of the logistic map may therefore be represented by a sequence of 1's and 0's, indicating its approximate position in these partitions at each iteration.

At $\mu = 2.9$, the system quickly descends onto a particular stable attractor state at approximately $x = 0.65$. The symbol sequence generated by this system is extremely simple – "1111…". The system in fact never leaves this interval, never passing the threshold into the other, and therefore emits the symbol "1" for any subsequent iteration once $F^n(x)$ reaches its attractor. However, when $\mu = 3.55$, for example, the map fluctuates for a bit and then reaches eight distinct and perfectly repeated attractors. Once it reaches this meta-stable state, while tracing the transitions across intervals I_1 and I_0, this

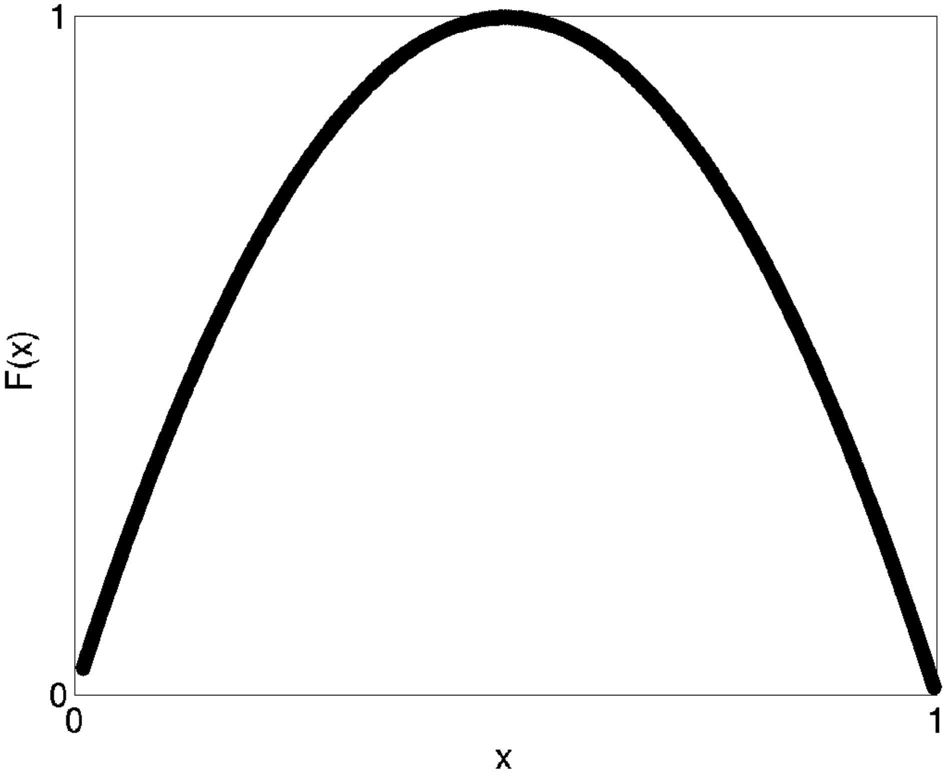

Figure 6.1: A phase-plot for the logistic map. Provided is above 0 and below 4, the system lives within the interval of [0, 1]. In other words, given its current value of x at time t, its subsequent iteration, time t + 1, will be on this line (here with μ = 3.9).

trajectory generates the sequence "011101110…". This may be simplified using the notation $(01110111)^n$ or even $(0111)^n$. Contained in this simple sequence rule is the original dynamics: Transitions between eight separate attractor states.

The above example is deliberately simplified for the sake of introducing rudimentary dynamical systems and symbolic dynamics. The logistic map affords this simplification. The strategy of employing symbolic dynamics,

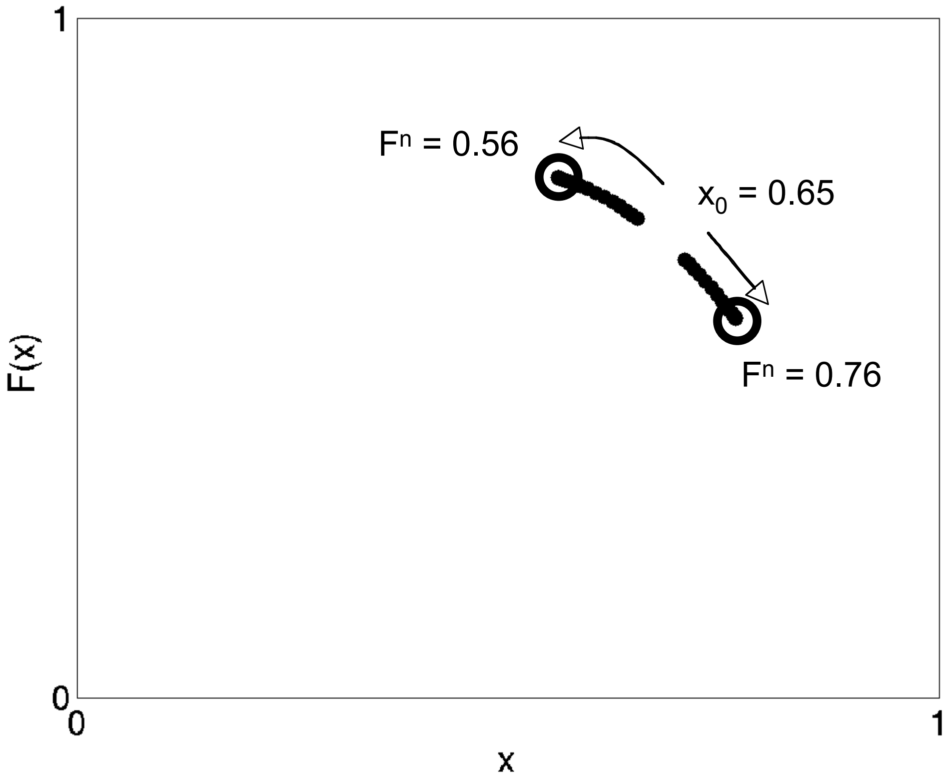

Figure 6.2: In this “phase flow” diagram, the logistic map moves into two stable attractors with μ = 3.1. The system starts at x_0 = .65. As the x is iterated through F, it settles into two attractor states, approximately .56 and .76, between which it will alternate indefinitely.

however, is somewhat more complex in most contexts. Symbolic dynamics rapidly served to help explore chaotic dynamical systems in more theoretical contexts (see Williams, 2002, for a review). In a further simplified example, we can straightforwardly introduce what this application entails. Consider an alphabet of N symbols that we might use in our partition of a system’s phase

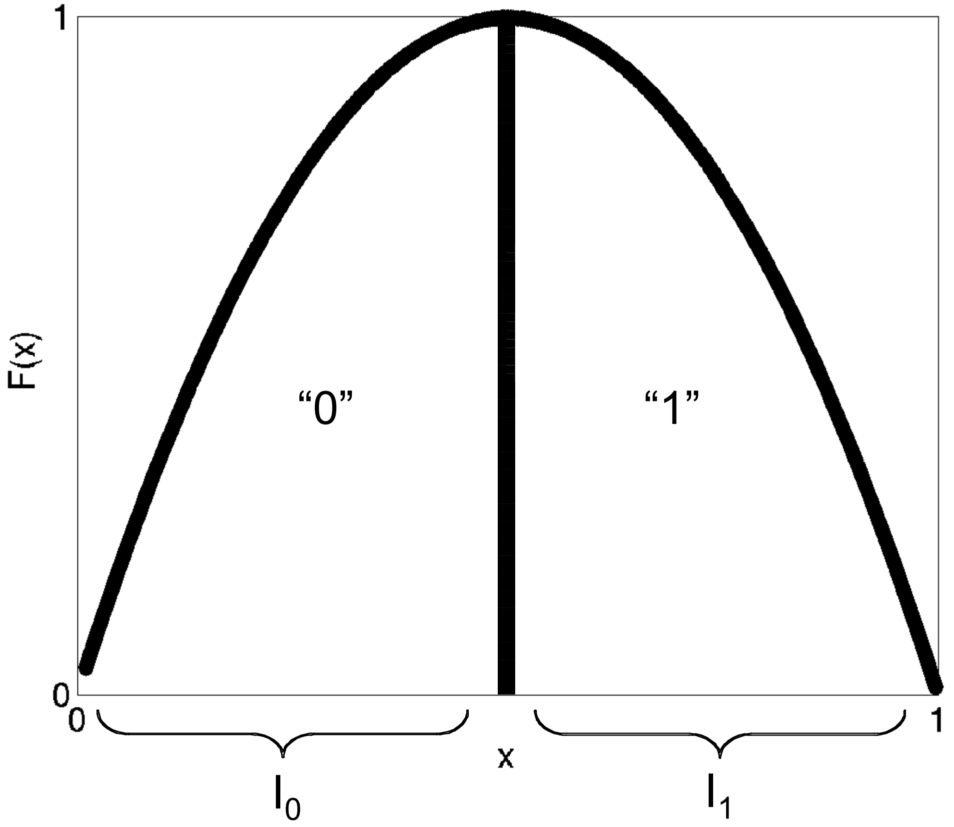

Figure 6.3: The phase space of the logistic map can be carved into two intervals. Each time the system enters the interval, it outputs the symbol corresponding to that region ("0" or "1"). Here $\mu = 3.9$ again.

space, A = {0, …, N-1}, and the space of all possible sequences constructed from this alphabet:

$\Sigma_2 = \{S \mid S = s_0 s_1 s_2 \ldots, \text{ and } s_i \in A\}$

Here, s_0 is the first symbol emitted by the dynamical system, and the sequence continues *ad infinitum*. The set Σ_2 is the space of all such possible sequences. A particular system's dynamics can be captured by shifting its infinite sequence, $S \in \Sigma_2$, to the left, so that $s_0s_1s_2\ldots \rightarrow s_1s_2s_3\ldots$, and the new sequence begins at the next emitted symbol, s_1. This shift operation captures the progression in time of emitted symbols, and is often represented by σ, so that $S' = \sigma S$, where $s'_i = s_{i+1}$. This shift operation can act as a mapping on a continuous space, $\sigma : \Sigma_2 \rightarrow \Sigma_2$, by specifying a distance measure or metric between sequences, $d(S, S')$. In other words, the trajectory of a system can be represented in terms of an ordered set of infinite symbol sequences, formed by progressive shifting.

From here, a means of exploring dynamical systems involves demonstrating that the space of symbols Σ_2 and its shift map σ have a certain geometrical equivalence to a dynamical system's own continuous mapping and the set of states which it visits. The famous Smale horseshoe can be studied through partitions of its phase space – and through its symbolic dynamics, it can be shown to have particular dynamical features (e.g., chaos; Medio & Lines, 2001). The logistic map has also been explored through its symbolic dynamics. Consider the case where the control parameter μ in F is larger than 4. It is easy to see that most initial states will have F^n approaching $-\infty$ as n gets larger. Specifically, since $x = .5$ grants the product $x(1 - x)$ its largest value (.25), any value for μ that is greater than 4 will take F outside the interval [0, 1], and thus at the next iteration, on a path towards infinity. However, inspection of the phase plot for the logistic map in Fig. 6.4 reveals the simple observation that not all values of x take F out of [0, 1]. The set C of all values that avoid this escape, along with the function F, can be shown to have this kind of equivalence with Σ_2 and σ, and allow certain conclusions about the properties of this set: Once again, the map F on C is chaotic (Devaney, 2003).

These textbook examples of the theoretical and mathematical benefits of symbolic dynamics merely scratch the surface of its recent role in the study of dynamical systems. Recent excitement has instead been concerned with the extent to which symbolic dynamics is informative about more complex systems through statistical analysis of its output. Symbolic dynamics is thus intriguing because it offers structures of sequences that can be subjected to a wide variety of "tricks for predicting discrete stochastic processes" (Shalizi, 2004a). Such statistical analysis has offered insight into complex dynamic processes in a wide variety of fields, including astronomy, biology, chemistry, and computational linguistics (see Daw, Finney & Tracy, 2003, for a review). The past two decades have also seen symbolic dynamics make explicit connections between the study of digital computation and that of continuous dynamical systems (Crutchfield, 1994).

Symbolic dynamics also has the interesting property of sometimes exhibiting equivalence with the continuous system from which it originates. As long as a partition is adequately selected, analysis of the symbol sequences can actually be used to reconstruct the continuous dynamics of the original system. A specific kind of partition, termed *generating partition*, can in fact yield "approximately complete and precise descriptions of the system" (beim Graben, 2004, p. 47). Perfect definition of a generating partition requires knowledge of the original dynamical system, but there exist techniques for approximating these demarcations (e.g., Davidchack, Lai, Bollt, & Dhamala, 1999; Kennel & Buhl, 2003). Such generating partitions allow the symbolic

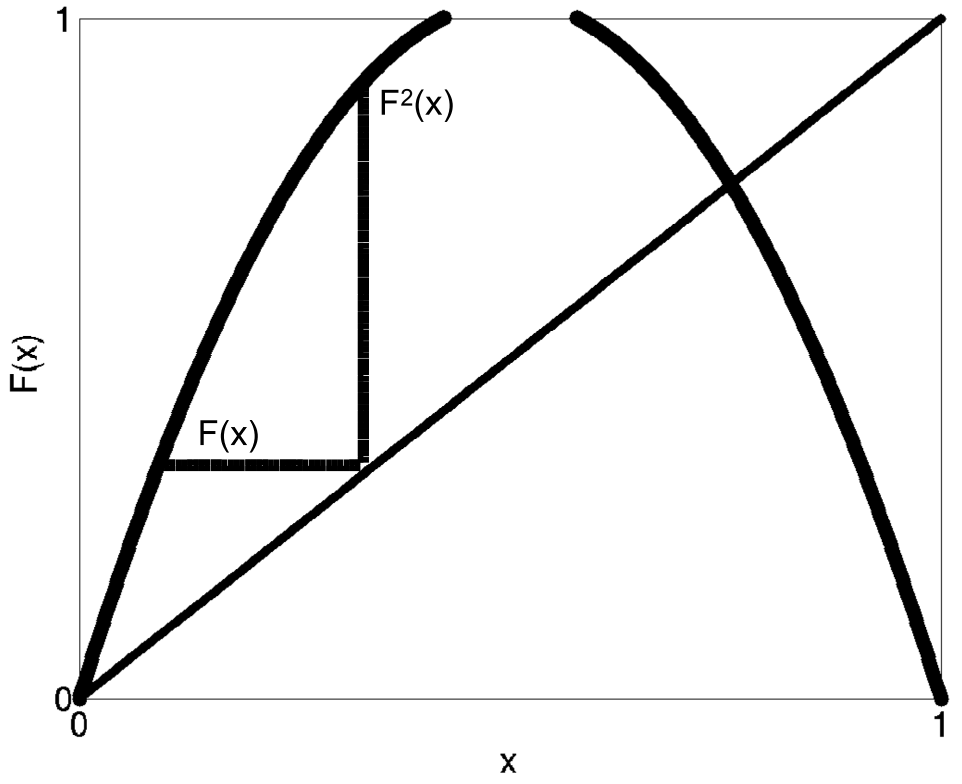

Figure 6.4: The logistic map phase plot with μ = 4.1. A portion of the phase space is outside the interval [0, 1], and points leaving will tend to -∞ through F iteration. However, a set of points does not leave this interval, illustrated with one iteration of some value x (dotted lines). The initial value x becomes F(x), and remains in the [0, 1] interval. Symbolic dynamics allows investigation into the nature of these iterations that do not escape (see text for more detail).

dynamics to be topologically equivalent to the original continuous dynamics (beim Graben, 2004; Kitchens, 1998; Shalizi & Albers, in press). However, finding generating partitions is very difficult in systems consisting of more than two dimensions (Kennel & Buhl, 2003), and they only work for deterministic dynamical systems (Crutchfield & Packard, 1982). Therefore, much of the practical applicability of symbolic dynamics may lie in iteratively refined approximations of generating partitions, rather than true generating partitions. For example, non-generating partitions in symbolic dynamics have been used for describing the phase-space of bimanual rhythmic coordination (Engbert, Scheffczyk, Krampe, Kurths, & Kliegl, 1998) and of heart rate variability (Kurths, Voss, Saparin, Witt, Kleiner & Wessel, 1995). However, with even slightly misplaced partitions, the threshold-crossing method for emitting symbol strings from continuous trajectories can very easily introduce severe compounded misrepresentations of the original continuous dynamics, i.e., grammatical errors in the symbol sequences (Bollt, Stanford, Lai, Zyczkowski, 2000, 2001). Symbolic sequences from more complex systems have therefore often been subject to more experimental or empirical styles of analysis (Daw et al., 2002).

To summarize, there have been two broad areas in which symbolic dynamics have made a clear impact. In the first, and ultimately its origin, it is explored extensively in pure theoretical contexts in mathematics to study tractable systems. In another, it has played a role in simplified descriptions and statistical analyses of more complex iterated mathematical processes, and even in application to dynamics of complex physical systems. It may, we argue, offer something to cognitive science theory as well. In the debate on representational format, symbolic dynamics could make headway toward formalizing theoretical debate. Current discussion on symbolic dynamics raises a number of important questions in this respect. We next consider these, and then introduce some reflections on future directions for symbolic dynamics in cognitive science.

4. Symbolic Dynamics and Cognitive Science

Dietrich and Markman (2003) actually describe something very close to symbolic dynamics in a short segment of their paper supporting discrete representations. They offer a number of properties that cognition must have, which only discrete representations endow. One of these properties is compositionality: Representations best explaining many cognitive processes must have component parts that are combined (see also Fodor & Pylyshyn, 1988; Marcus, 2001). They argue that any representational subsystem, if continuous, can only have parts if there is some other system that discretely interprets its regions, and takes in discrete representations as input. This is in fact a description of symbolic dynamics, though there are details to be worked out. For one, the resultant symbols, if not time-course irrelevant, might encode the original dynamics of the system, as mentioned in the way of generating partitions above. This would mean an equivalence relation holds between the two systems, at least in the sense that the symbols carry some of the continuous information in the original dynamic subsystem. Secondly, it has been demonstrated recently that dynamical systems actually do have considerably surprising computational powers. In fact, a number of these properties often considered hallmarks of discrete-symbolic algorithmic processing can be approached with symbolic dynamics.

For example, one such property, *discrimination,* is easy to achieve through translation into symbol sequences. Multi-stable one-dimensional dynamical systems can emit symbols pertaining to any stable point (and a given interval around it) in its phase space – as a matter of discrete, symbolic output from that system. This scenario may indeed be superior to verbalized discrete theories since symbolic output from an iterated map retains some information about time. For example, a meta-stable system that drifts slowly will produce symbol sequences with long strings of identical symbols, indicating its inhabiting of some categorical state. The output is therefore discretely representational, but also reveals patterns of change in time. Perceptual state-space is of course not a matter of collapsing over a single dimension – the situation becomes very complex when we consider the number of categories (symbols) that need to be represented, and the fact that our visual system, for example, is translation invariant, so whatever partition can define the symbolic output from visual to cognitive processes must exist in a very large number of dimensions. A second problem concerns delineating the stages at which such collapsing from continuous mappings to discrete symbol strings occurs. As mentioned, to account for continuous perceptual states moving into something akin to sets of sameness, there must be one or more stages of "analog-to-digital" (A-to-D) conversion.

Probably the most studied and established property that dynamical systems exhibit through their symbolic dynamics is their digital computation – that a description of dynamical systems can take the form of explicating its information-processing capacities via symbol sequences. This feature approaches the well-known issue of *systematicity*, a property that many have argued cognitive systems must have (especially human ones; e.g., Fodor & Pylyshyn, 1988; Hadley, 1994; Marcus, 2001). This discussion fits into the scope of symbolic dynamics in three ways. First, it is possible to characterize the dynamics of a system through computational descriptive schemes. For example, the classic paper by Crutchfield and Young (1989) introduced an approach to nonlinear dynamical systems that quantifies their computational qualities. Subsequent research has pursued the extraction of such intrinsic computation from nonlinear dynamical systems, among other systems (see Andrews, 2002; Crutchfield, 1994, for reviews). Sought after qualities of systematic computation urged by Hadley (1994) and Fodor and Pylyshyn (1988) may very well be encoded in the edge-of-chaos dynamics of even simple systems (Crutchfield & Young, 1990).

The second way systematicity can fit into symbolic dynamics is through exploring the ability of dynamic systems to acquire formal languages. The system that learns the language may again be characterized in terms of symbolic dynamics. For example, the well-known simulations in Pollack (1991) demonstrated that a neural network can learn context-free languages and classify novel sentences from such grammars via a decision process akin to symbolic dynamics (see also Cleeremans et al., 1989 for related examples). More recently, Rodriguez, Wiles and Elman (1999) investigated a very simple recurrent network in its ability to learn deterministic context-free grammars. Networks that learned successfully performed a form of "counting" in their phase space. This allows successful learning of the context-free grammar without explicitly implementing a pushdown automaton. Also, Tabor (2001) recently used a neural network model trained to predict sequences of symbols from four languages of differing levels of complexity (see also Tabor, 2000). Networks trained on context-free languages, as opposed to the regular languages, exhibited edge-of-chaos effects (or intermittency), revealing the kind of intrinsic computational qualities outlined in other nonlinear dynamical systems by Crutchfield and colleagues (Crutchfield & Young, 1989).

Finally, one may simply take the symbols for granted at some level. In this case, though the solution appears simplistic, the difficulty is in delimiting and discovering the nature of the interface between continuous perceptual states and the resultant discrete cognitive informational states that undergo algorithmic manipulation. Resorting to this idealization, and thereby taking for granted straightforward algorithmic parlance about representations, requires explication of A-to-D conversions between high-dimensional continuous perceptual states and their entryway into cognitive processing. Although the above review suggests the continuous dynamics of simple systems can already exhibit surprising computational qualities, it may be in the domain of this symbolic demarcation that debate between formats is best mitigated. Following our discussion here of theoretical issues in the use of symbolic dynamics, we offer some sketches of the ways in which this mitigation might take place in computational models.

There are numerous issues with both symbolic dynamic theory and application that are relevant to theoretical frameworks in cognitive science. We introduce three of these issues here, and elaborate further in the subsequent sections of this article. The first issue concerns the consequences of generating partitions. Some have argued that the equivalence between symbolic dynamics (from a generating partition) and the continuous mapping from which it originated, renders moot the debate between continuous and discrete states in the mind (Crutchfield, 1998; Shalizi, 2004b). Though compelling at first pass, the argument is based on simple, low-dimensional systems – ones whose consequences cannot be handily generalized to noisy, high-dimensional (and likely highly non-stationary) dynamics in neural systems at the level of cognitive processing. As a second issue, we consider the implications of recent discussion concerning the epistemic problems of finding appropriate partitions for meaningful symbolic dynamics (beim Graben, 2004). This has direct relevance to conceptualization of “error” in competence and performance, and the nature of language comprehension and production, among other cognitive processes. The final issue concerns how continuous and symbolic dynamical systems function in tandem during perception and cognition. Presumably, if discrete-symbolic descriptions are most suitable for “higher” cognition, then there must be some stage at which continuous dynamics of perceptual or “lower” cognition gets transmogrified into interpretable symbolic states.

4.1 Continuous-Symbolic Equivalence

As discussed above, there are many reasons to study continuous dynamic maps via symbol sequences. An interesting fact for many such maps is that there exist generating partitions that emit symbol sequences exactly reflecting the original dynamics of the system. This has led some to argue that the debate concerning discrete-symbolic and continuous-distributed representations is ill-posed (e.g., Shalizi, 2004b). Since a dynamical system can be seen as identical with some symbolic dynamics, it might be inappropriate to suppose that two formats of representation are at odds when they are mathematically equivalent. As already mentioned, Crutchfield and collaborators (Crutchfield, 1994) conceive of dynamics as inherently computational, and offer numerous techniques for generating computational machinery from symbolic sampling of continuous states (e.g., ϵ-machines; Crutchfield, 1994). Elsewhere, Crutchfield has argued that supposing dynamics can replace discrete computation (e.g., Van Gelder, 1998) neglects the intrinsic computational nature of dynamical systems themselves (Crutchfield, 1998).

Early in Crutchfield's seminal paper (1994), he distinguishes between two concepts of computation. The first, "useful" computation, refers to specific instantiations of input-output mappings in some computational architecture. The second, "intrinsic" computation, concerns the basic capacities and limitations of a computational system, dynamical or otherwise. This involves exploration or specification of information-processing capacities of a system, without reference to any specific "useful" input-output accomplishment. This perspective has led to extremely fruitful research on discovering the underlying computational aspects of nonlinear dynamical systems. For example, early work by Crutchfield and Young (1989) sought to specify and measure the complexity of minimal stochastic automata whose state transitions (emitting symbols) embody the logistic map's dynamics at differing values of μ (see also Crutchfield & Young, 1990).

There are, however, reasons for remaining cautious about the direct implications in our understanding of complex cognitive states. An in-principle statement concerning the equivalence of continuous and symbolic dynamics in systems is not sufficient on its own to alleviate debate. There are at least two related reasons for this. First, relying on such equivalence neglects the very crucial and substantial details that debate on representational format carries. A rather straightforward one is the kind of characterization that symbolic and distributed formats receive. For example, Andy Clark (2001) characterizes much symbolic cognitive theory as resting on representations whose contents are *semantically transparent*. A classical computational theory of language deals in representations of words, their meanings, and the structures that they compose. These representational formats are highly "scrutable," their significance in a system's computation immediately accessible. However, systems relying on probabilistic and distributed representations or states, such as connectionist systems, often rely on formats that are *semantically opaque*. For example, establishing the function of a hidden-unit manifold often involves detailed statistical analysis of the hidden-unit activations under varying circumstances. The resultant function may be very nonlinear and complex, and not easily describable through commonsensical or folk-psychological labels.

For this reason, simply saying that the two kinds of descriptive machinery, continuous and symbolic, both serve the same functions actually skirts some substantive issues. The debate concerns explanation in terms of specific kinds of computational mechanics – concretely identifiable words in our "language of thought" (Fodor, 1983), or some other more or less semantically transparent discrete states. These are pitted against models accounting for behavior in terms of distributed representations whose interpretations are less obvious, or perhaps "subsymbolic" (Smolensky, 1988). In fact, cognitive science has already had a number of battles concerning whether these two systems are equivalent, or the second being just a special case of the first, and so on (e.g., Fodor & Pylyshyn, 1988; Lachter & Bever, 1988).

The second reason to be cautious about the lesson from symbolic-dynamic equivalence is that "useful" computation has been considerably less explored than "intrinsic" dynamics in the study of computational mechanics. Although the current accomplishments can only be described as some of the most exciting and relevant to cognitive science, they have yet to delve into systems whose complexity can match a level of description needed for understanding *cognitive* processes. Van Gelder (1998) replies in this manner, remarking that when "it comes time to model the complexities of real cognition—to publish in Psychological Review rather than Physica D—they may find that the dynamics drops out of the picture, and the relevant story is cast entirely at the level of the emergent computation. Alternatively, they may find (as have many dynamicists) that the computational aspects play second fiddle to the dynamics" (p. 13). This seems to misunderstand what is accomplished in symbolic analyses of dynamical systems: The descriptions are two sides of the same computational coin. A more direct concern at present is whether meaningful partitions can be established. As we approach a level of complexity that matches what is accomplished in a neural substrate, or proposed cognitive processes of multiple dimensions, the likelihood of finding generating partitions drops radically (see next section).

So, whether or not we embrace the equivalence of symbolic and continuous dynamics through generating partitions, we are still left with some confusion. Are the discrete symbolic states of our cognitive system available for scientific scrutiny, and the dynamics more complex (yet equivalent)? Or are discrete symbolic states of our cognitive system inadequate explanatory constructs, and we should reach for continuous dynamic descriptions of our mind/brain? Churchland (1992) offers discussion relevant to these more substantive issues in the domain of neural networks, and considers partitions of their hidden-unit state space that can reflect conceptual structure in human cognition. Adopting a set of partitions, Churchland argues, "may suffice for the accurate short-term prediction of its behavior, but that knowledge is inadequate to predict or explain the evolution of those partitions over the course of time" (p. 178). We argue that, in the domain of higher levels of cognitive processing, this position has considerable merit, but is very much without consensus in the field. It is thus through these substantive issues that the two formats of representation stake their respective claims.

4.2 Epistemic Issues

Similar to Andy Clark's (2001) distinction between transparent symbols and opaque distributed representations, Atmanspacher (2000) makes a distinction between epistemic and ontic types of description of chaotic systems. An ontic description is exhaustive concerning the dynamical system – it comprehensively encapsulates the composition of the dynamics. An epistemic description is framed in terms of knowledge or ignorance of an observer evaluating these ontic "states." Epistemic descriptions are achieved by evaluation of an observed or measured dynamical system, through statistical quantification or characterization of it. This terminology of *ontic* and *epistemic* can be used to frame the previous section's discussion of equivalence through a generating partition. The pure equivalence between a symbolic dynamics and its origin map can involve only ontic descriptions for any system of sufficient complexity. That is, urging the equivalence of some symbolic dynamics and the original system implies a kind of ontic state that is inaccessible to us as observers. Instead, we are confined to epistemic descriptions for complex cognitive systems. In order for a chosen symbolic dynamics to adequately explain or represent the cognitive process under study, it must be chosen appropriately. This is not a trivial matter. We consider two problems of these epistemic issues relating to finding a good partition for a continuous map.

First, beim Graben (2004; beim Graben & Altmanspacher, 2004) argues that incompatible, but equally accurate, symbolic epistemic descriptions are possible with multiple non-generating partitions. This means that two sets of different symbolic dynamics may be equally adequate as formal descriptions of the original dynamics, yet mutually incompatible with one another (cf. Quinean indeterminacy: Quine, 1960). beim Graben (2004) provides an example of a Hopfield network as a multi-stable dynamical system living in a space of many dimensions (Balkenius & Gärdenfors, 1991). There is no generating partition for this space, and indeed multiple descriptions via symbolic dynamics can be mutually incompatible, while remaining equally good (or bad) partitions of the underlying dynamical system. As a consequence, non-generating partitions can provide "conceptual" descriptions of the continuous system, and there may indeed be many such descriptions. While all of them may serve as formal descriptions at a symbolic-discrete level, they can be mutually incompatible with each other. Quine (1960) early on made very similar points concerning the study of linguistic meaning by indicating that many equally good (or bad) rule sets can exist for translation from one language to another; yet these equivalent translation strategies may be mutually incompatible when compared directly (see also Moore, 1956, for a related classic theorem; Gauker, 2003). It is crucial to point out that this is not merely a fact of further exploring the ontic description of the given system in order to select the better of these incompatible symbolic accounts. In fact, Quine, and beim Graben and colleagues in symbolic dynamics, reveal that even given the full set of ontic states themselves, there are still mutually incompatible, yet thoroughly equally good, symbolic descriptions. In short, they cannot be reliably adjudicated among.

A second issue concerns the kinds of errors that result from inappropriate partitions. Bollt et al. (2001) analyze the tent map, whose generating partition is known, and measure the topological entropy resulting from shifts of that partition to varying degrees. The effects of shifted partitions are quite drastic, with topological entropy being affected immediately, and in an irregular (non-monotonic) fashion. The upshot, according to the authors, is that arbitrary partitions (e.g., Kurths et al., 1995) can result in “severe under-representation of a dynamical system” (p. 281). It is important to point out that these results were based on a well-known deterministic and simple dynamical system. The effects of noise (Crutchfield & Packard, 1982) and increasing complexity (Kennel & Buhl, 2003) in degrading the fidelity of the emitted symbolics are also well documented.

The foregoing remarks on epistemic limitations on symbolic dynamics have two implications for current discussion. First, they serve to underscore the points made in the previous section on equivalence. The possibility of the existence of a generating partition is not sufficient to dissolve debate on symbolic versus continuous representation. Instead, the epistemological limitations on more complex, noisy dynamical systems suggests that there is considerable room for debate concerning the adequacy of either continuous or symbolic accounts for some cognitive process. Indeed, in most cases (if not all), we do not have sufficient knowledge of the ontic conditions of some cognitive process. The upshot for cognitive science is that continuous or symbolic accounts are 1) highly unlikely to be resolved by mere recognition of equivalence, and 2) are likely to offer differing amounts of coverage of the human data regarding these ontic states – about which the field has much to discover.

4.3 A-to-D Conversions

Despite these limitations, we argue that the promise of symbolic dynamics lies in articulating the transition from dynamical, continuous descriptions of perception, into the theoretical language of discrete, algorithmic processes for high-level cognition. Whatever the *ontic* states underlying cognition, our *epistemic* descriptions and theories ought to be couched in structures or processes that bear causal relationships to others, and ultimately, to our observable behavior. If it is the case that some are discrete and symbolic, there must occur a transition into them from a continuous state-space of perceptual or early-cognitive processing. These A-to-D conversions consist of collapsing the continuous-distributed representations onto discrete-symbolic ones that cause inherent information loss about the perceptual states feeding into them. However, this loss may be merely "lossy," as in image compression algorithms, where the resulting compact representations still carry information appropriate for cognitive processing.

The question is then not merely *when* this transmogrification occurs, but also *what kind* of information from continuous states do these discrete states need in order to account fully for observable behavior. For example, in information-processing accounts of cognition, Miller's (1982) exploration of the concept of information "grain" provided an early challenge to discover what kind of discrete representations there are: What level of "granularity" do discrete representations need to have to account for cognition. For example, along with Miller's work on the "response preparation effect," Abrams and Balota (1991) and Spivey, Grosjean, and Knoblich (2005) demonstrate that dynamic response measures (e.g., force and velocity measures from a response bar or continuously recorded computer-mouse movements) exhibit graded properties depending on the continuous strength or reliability of the information that produced the motor output. In addition, a wide variety of eye-movement research (reviewed above) suggests that metabolically cheap movements such as saccades reveal a decision process that does not appear perfectly discrete. These studies demonstrate that "echoes" of continuous information states can be observed in the *dynamic properties of resultant responses*. The discrete states that may have mediated the transitions from sensors to effectors must carry at least some relevant information from early graded states. In other words, while reaction time may reveal information about the decision process during discrete, algorithmic processing, the concomitantly graded manual output from the system indicates that even when these discrete decision processes collapse onto the effectors, there remains some fine granularity.

We can frame the situation simply by defining an idealized "problem space" for some cognitive process. The space may be maximally simple, from an idealized or simulated continuous perceptual space into one or two symbolic processes. Here, A-to-D conversion performs a *useful* computation in the Crutchfield sense, described above. Whatever the intrinsic computational properties of the initial continuous perceptual space, the system of continuous representation feeding into discrete symbolic processing has an informational function in a problem space of, say, evolutionary relevance (for some such thing as reproduction, or running away from something that might eat you).

A simple example is perceptual categorization. An idealized continuous space of perceptual information can be manipulated so as to output symbols feeding into some discrete process. This idealized scheme may be suited for existence proofs of granularity of the A-to-D conversion to adequately account for such graded effects outlined above (see Churchland, 1992, for some early possible examples).

Symbolic dynamics offers a playground in which this conceptual problem can be formally explored. Given a dynamical system living in a state space of m dimensions, a set of stable or meta-stable attractors can be explored via simulation. Like perceptual processes, this dynamical system can feed into a separate system, described in a variety of ways (e.g., a Turing machine, or finite-state machine), that receives symbolic input via threshold-crossing in a partition carving that state space. This collapse involves maximal loss: From m dimensions into 1 dimension of Q possible states defining the partition. Of course, for simple systems such as the logistic map (where $m = 1$), this collapse can still carry the original dynamics, and entail some interesting computational properties. As for higher-dimensional state spaces, there seem to be two ways that models of this kind might begin to approach existence proofs for symbolization of a perceptual space. On the simplest side, one might explore resultant algorithmic processing on symbolic output of just one dimension with Q regions of one partitioning. These states may be numerous, or *refined*, enough to carry some echoes of the original space. Another possibility is to consider collapsing the state space of m dimensions into more than one partition. The m-dimensional system may be collapsed onto an n-tuple of symbols, each element of which is the output from some separate partition that uniquely carves the state space. Exploration of this system would involve subsequent A-to-D/D-to-D conversions, permitting sequenced levels of granularity in the various stages of perceptual-cognitive processing (that is, without considering feedback projections).

There exist a number of “useful” computational models performing symbolization of this sort. For example, the decision processes of Pollack’s (1990) dynamical recognizer and the Hopfield network of Balkenius and Gärdenfors (1991) that implements non-monotonic logics are two relatively early models. Recently, Tabor has specifically addressed the learning and processing of formal languages by such systems (Tabor, 2000, 2001), along with beim Graben and colleagues’ sophisticated analyses (beim Graben, Jurish, Saddy, & Frisch, 2004). These are just a few of the enticing invitations to employing symbolic dynamics in a way we recommend here: Devising existence proofs relating the stages of continuous-discrete transitions in a simplified problem space akin to cognitive processes and their output. Though only promissory at this point, symbolic dynamics may make it possible to reconcile both the dynamic and discrete descriptions of the states and processes underlying cognition.

5. Conclusion

We do not have full privileges in our access to the ontic states of our mind/brain. An inevitable fact about higher-order cognitive theories is that they are descriptive at a very coarse level – it is currently an intractable problem to specify, even partially, the dynamics underlying neural computation in a cognitive task of any nontrivial complexity (e.g., Uttal, 2001). Nevertheless, for the majority of cognitive scientists, this daunting state of affairs does not invalidate proposals for structures and processes of our cognitive system. In this article, we have limned the surface of a framework within which competing theoretical accounts of representational structures and processes may have equal opportunity to contribute to our understanding of cognition. Symbolic dynamic investigations of idealized problem spaces may provide a common arena for exploring the interplay between continuous-distributed and discrete-symbolic representational accounts. Moreover, as a framework for further discussion, it may help both representational formats overcome the limitations of time-course irrelevant descriptions of cognition. Given a set of input information and informational goals, symbolic dynamics offers both informational and temporal insight into the transition from continuous perceptual trajectories into more or less fine-grained discretized states for higher cognitive processes. There will, of course, be some conceptual and technical obstacles in the way ahead, and we have considered a number of these above.

Given the deep epistemic problems experienced by all theories of cognition, and the complexity of the brain onto which they are imposed, it seems we might forever be confined to epistemic descriptions of the ontic states of our cognitive system. Regarding the state of the art in cognitive science, the dispute between two such families of description, discrete-symbolic and distributed-continuous, thus seems just as likely to be evenly conciliated than for one or the other to win permanent prominence. This article offers some further considerations of symbolic dynamics to contribute to ongoing debate (see also, e.g., beim Graben, 2004; Goertzel, 1998). Mathematization of simplified problem spaces, such as perceptual categorization or computation in "chaotic itinerancy" (Tsuda, 2001), may be the route to a formal terrain permitting cohabitation of both kinds of theoretical constructs – or, at least, a mutually supportive arena in which they can have a fair fight.

It is perhaps a striking illusion, at the physical level, that there exist discrete states of the mind/brain. This at least seems to be the case if you grant that the substrate is in constant motion, like Heraclitus's river. The illusion is nevertheless difficult to overcome, because our phenomenology seems to be in an inescapable embrace with experiences that have strict boundaries. At the epistemic level, rather than the phenomenological level, it may be inevitable that boundaries need to be placed around neurophysiological complexity to construct sufficiently explanatory, and tractable, theories of cognitive processes. This becomes evermore troublesome when one considers functional redundancy and feedback loops within the substrate, as well as between it and its environment. So while theoretical debate may continue concerning whether the mind is a system that imposes boundaries on a continuous information flow, symbolic dynamics may offer a mathematical terrain in which these boundaries can be rigorously explored.

CHAPTER SEVEN

Discussion: Matters of Fact and Units of Analysis

Resummary

Chapter 1 presented the history of the preponderance of the symbolic computing metaphor in theories of cognition. This metaphor provided justification for separating cognition from action, thus relegating response dynamics to a lesser role in understanding intelligent behavior. The evolutionary thought experiment of Braitenberg recommended a different position on this issue: Thinking of cognition as increasingly complicated, but inherent, perception-action mediation. This strong interpretation of the thought experiment suggests that response dynamics should reveal "echoes" of cognitive processing. The prediction was made that even in a higher-order process such as categorization, response dynamics might reveal aspects of cognitive processing.

Chapters 2 to 5 substantiate this prediction. Chapter 2 revealed that in both lexical and perceptual modes, animals of high atypicality drew the manual responses towards a competing category, particularly if that category had overlapping features. Chapter 3 motivated the modeling endeavors in Chapter 4, which demonstrated that the response dynamics observed in Chapter 2 could be generated by a system in which processing and responding interact bidirectionally. A second simulation predicted finer-grained effects observed in Chapter 5. Typicality gradients are reflected in response dynamics, and in fact, feature processing specific to lexical categorization was reflected in motor output.

Chapter 6 urged a mathematization of representation that is independent of specific models and architecture currently under dispute. It was argued that symbolic dynamics can serve as an arena in which the frameworks of the discrete computer and continuous dynamical metaphors might be more evenly compared. Extensive discussion in that chapter argued that substantive issues in cognitive science could be interestingly explored using such a common descriptive system as symbolic dynamics.

The current, and concluding, chapter addresses what may come of this kind of theoretical debate from a more general philosophical perspective. The symbolic dynamics approach seeks to adjudicate between theories vying for the matters of fact of cognition. As discussed in the Chapter 1, frameworks in cognitive science aim to account for representation and processing as it is in reality -- the "end-of-discussion" characterization of our mental inner-workings. The next section discusses the potential outcome of such an aim. Are cognitive matters of fact attainable, and of what sort? The final section of this chapter presents a perspective on explanatory frameworks that is plurastic and pragmatic in nature. It is contended that the goals of cognitive explanation are in fact not to detail the matter-of-fact nature of the cognitive system -- but rather, to solve specific problems within particular substantive domains. The upshot: Explanatory frameworks of varying natures have varying forces of explanation and prediction. Those specific research problems are amenable to these forces of explanation on a case-by-case basis. When we have successfully solved numerous problems, then it becomes interesting to bridge these explanatory frameworks. It is here that such metatheoretical approaches like symbolic dynamics might not just "mitigate," but *integrate*.

Chapter 6 argues for the use of a neutral mathematization like symbolic dynamics to discern the value of discrete versus continuous explanatory frameworks. Symbolic dynamics, and other systems permitting comparison of one kind of representation versus another (e.g., hybrid systems; e.g., Sun, 1997), are necessary given the inevitable limitations in the epistemic access to the ontic states of the cognitive system (see Chapter 6).

Over time, "in the limit," use of such descriptive machinery as symbolic dynamics may give way to a few possible theoretical outcomes. The simplest is that one framework will exhibit comprehensively superior accounts for the data. For example, the continuous-distributed framework may become the most comprehensive framework within the competing arena of symbolic dynamics. Such an outcome is a naively hopeful one. First, "in the limit" above must inevitably mean "after a very long period of empirical research," because the current available data, in the various subdomains of cognitive scientific inquiry, recommend only a tempered prediction about which framework might become comprehensive. As mentioned in Chapter 6, these two broad frameworks enjoy differing levels of success, depending on the subdomain you consider. This will become particularly relevant below.

Another problem with such a simple theoretical outcome is the extremely hopeful assumption that "in the limit" cannot mean "after an infinite period of empirical research" -- that all evidential eras will involve competition between two competing frameworks. It is an assumption that a particular (and likely lengthy) period of time will pass after which comprehensive data coverage will be achieved. The history and philosophy of science suggest this not to be the case, though there is more detail to this point, further considered below.

Sorting out the matters of fact about the cognitive system will likely remain an active debate for some time, and investing in the one-winner outcome is merely hopeful at present. Nevertheless, if one were pressed to decide on matters of fact, given the current (or any future) evidential situation, there seem to be three prominent possibilities. These three philosophical positions in deciding matters of fact, the correct "cognitive ontology," are presented below. The first, a kind of *Quinean confusion*, is that matters of fact, given any set of data, will never be determined successfully. The second, through *Cartwright's reality toolbox*, is that matters of fact are decided within each self-contained subdomain of cognitive science. The final option is that consideration of low-level processes (upon which higher levels are based) motivates the selection of one framework over another, particularly in view of the kind of evidence presented in Chapters 2-5 -- *Spivey's echoes*.

Matter of fact #1: Quinean confusion

Imagine you could survey the set of *all behavioral data* accounted for by cognitive frameworks. These data would include facts about memory performance, recognizing words versus non-words, how attention is focused in experimental contexts, and so on. Each datum ought to be predicted and explained by a successful framework (perhaps in the way that either mechanism- or process-based explanations work; see Bechtel, 1998). Multiple frameworks in opposition offer different sets of interrelated theoretical structures and operations that generate predictions about behavioral data. For example, the symbolic-discrete framework often offers up discrete rules and representations that frame the cognitive system in terms of a digital computer metaphor discussed extensively in the introduction. In this sense, the theoretical framework is giving us a "translation" of the behavioral observations into deeper theoretical commitments. These commitments organize the observations, and generate expectations about future ones.

Imagine this set of behavioral data is expressed purely in terms of observables. For example, one datum may be, "Participants respond significantly faster to words that have higher frequency."[1] The symbolic-discrete translation of these data may look something like the following: "Memory for words is organized as a list of discrete lexical entries, ordered by frequency. The mind accesses this list by order when recognizing a word," or something of that sort. "Memory" and "mind" and "access" provide the language of theoretical constructs and processes into which the prosaic descriptions of the data will be translated.

There are two important perspectives from the philosophy of science that apply to this situation. The first, a near consensus, is referred to as underdetermination of theory by evidence. The second, more controversial and from Quine (1960, 1970), is dubbed indeterminacy of translation.

Underdetermination of theory is the perspective that, given any amount of data, there will always be one or more other (possibly incompatible) theories that also satisfy these data. This does not entail that

---

[1] I do not consider the problem of "theory-ladenness" regarding such observational statements. It should be noted that the theory which lades the behavioral data may be one of a folk psychological nature, or some other scheme of current cultural forces guiding English descriptions.

there is no matter of fact about which of the theories is correct; it is an epistemological problem, in that the structure of scientific theories and the limitations of evidence give way to a continuous refinement and flexibility of individual theories to fit that evidence. Theories have very comfortable wiggle room to account for potentially confuting findings. This may be particularly true of theories that do not enjoy full mathematization (Meehl, 1998).

Given our set of behavioral data, and some tentatively adopted theoretical translation, there will be one or more further cognitive frameworks that account for the same data, but offer constructs and processes that are not mutually compatible. Often, certain theoretical virtues or constraints are thought to apply to these frameworks, such as simplicity, comprehensiveness, novelty of predictions, and so forth. For example, van Gelder and Port (1995) feel that, all things being equal, dynamical theories serve us better than connectionist or symbolic accounts because they involve a continuous-time process, thus being more intuitive, and often use fewer degrees of freedom, thus being simpler. Intuitiveness and simplicity are two such virtues or constraints on theory selection. The way these constraints apply turns out not to be very obvious -- for example, a fairly ornate, complex theory will be selected over others if it admits of the most intuitive means of translating behavioral data. These virtues are not a guaranteed guide to overcome these problems (Putnam, 1973).

Indeterminacy, on the other hand, is a much stronger point about this theoretical translation of behavioral data. If one accepts the view that cognitive theories offer a theoretical language into which behavioral observations are translated, then this situation invites application of Quine's so-called "argument from above" for the indeterminacy of translation (Quine, 1970).

This argument leads to indeterminacy rather than underdetermination, a much stronger thesis: Not only is there underdetermination of scientific theory in our scenario, but in fact, there is *no matter of fact* about which of the competing frameworks is correct -- given any and *all possible* behavioral data, one cannot adjudicate between competing theories because, even though they are not mutually compatible, they are "empirically equivalent." Indeterminacy is thus the stronger statement that one cannot decide on which explanatory framework is the correct one, given any and all evidential scenarios.

The "argument from above" works like this. Imagine translating a foreign physicist's theory into your own physical theory expressed in English. Both theories are underdetermined by available physical evidence -- but this underdetermination is compounded by the flexibility of translating one language to another. In other words, the underdetermination is exacerbated by a secondary underdetermination of translating the foreign language into English (see Quine, 1970; Miller, 1998, for an elaboration and some issues with the argument). In other words, we have the first case that the foreign physicist's theory may be translated into, say, M other possible accounts for the set of data (underdetermination #1), but it cannot be certain when translating this theory into another language such as English that the theory translates into the same original account out of those accounts available (underdetermination #2).

In our case, most philosophers would not see our cognitive framework translation as anything more than the standard underdetermination problem (e.g., Chomsky, 1968; Miller, 1998). However, details in Chapter 6 may be echoed here. There may be an in-principle equivalence between computation-based and dynamics-based epistemic accounts for an ontic system. Given any and all data obtained through measurement of this ontic system, which of these two explanatory frameworks is “correct” is irrelevant at best, or a false question at worst: There is no “correctness” attached to one over the other, because both capture the data adequately. However, while the systems are mathematically equivalent at root, their peripheral ontological commitments give “intuitively incompatible” systems. In other words, carving a state space into discrete partitions may motivate labelable regions that are not obvious if we maintained the continuous, dynamical descriptions. If we accept this possibility, then the situation is in fact not just underdetermined, but indeterminate.

This indeterminacy argument is less compelling than straightforward underdetermination. But in either case, the situation regarding matters of cognitive fact are troublesome: Given all our behavioral data, there will likely always be theoretical dispute. There may be a matter of fact, a “correct cognitive ontology,” but it will not obtain.

Matter(s) of fact #2: Cartwright’s reality toolbox

Quinean confusion assumes that the matters of fact of nature are obtained by adopting one explanatory framework. Only when one theory reigns over the competition can it be said that we have discerned the way nature in fact is. There is a surprising, but strong, motivation to reject this fundamentalist assumption. One prominent example is Cartwright (1999), who presents arguments that the nature of the universe is multifarious -- given to a variety of “natures” and “capacities” that are not confined to one theoretical system, but in fact many, and those many natures and capacities are discovered in varying contexts and subject matter. Each such system has rightful claim over the matters of fact of its relevant domain.

Cartwright (1999) contends that science gives us a patchwork of theories and laws that uncover the *natures* of the universe. Her detailed position is based on some analysis of current theories in physics and economics, and a novel conceptual contribution to philosophy of science. First, she notes that even the most successful fundamental physical theories cannot help us in situations whose conditions are drastically detached from those which commonly support those fundamental theories. For example, consider letting a dollar bill get taken by the wind in an open area. Fundamental laws of mechanics cannot predict this complex scenario -- it is of "limited serviceability" (p. 27). Rather, the problem may get shunted into the realm of fluid dynamics, which may "provide a practicable model." (p. 27)

Cartwright offers some novel conceptual devices for understanding how science works. She sees science's ontological toolbox as scientific experimentation -- physical models which arrange for measurement or other observation. According to Cartwright, scientific experimentation rigidly arranges parts of the world, and the natures of these parts are discoverable through these arrangements. Cartwright gives the name *nomological machine* to this perennial setup, and "all things being equal," it is a system for reliably generating certain outcomes given its conditions. Laws, and other regularities, are discovered through nomological machines. Nomological machines are thus the practice of designing rigidly controlled worldly circumstances, and accounting for the regularity seen therein.

The nature of the universe is thus not unitary. *Natures, plural*, are what it instead exhibits. Cartwright shares two different fables of God and Saint Peter to describe how a “metaphysical pluralist” might conceive of God’s role in creating our universe:

> In the first story, God is very interested in physics. He carefully writes out all of its laws and lays down the initial positions and velocities of all the atoms in the universe. He then leaves to Saint Peter the tedious but intellectually trivial job of calculating all future happenings, including what, if any, macroscopic properties and macroscopic laws will emerge. That is the story of reductionism. Metaphysical pluralism supposes that God is instead very concerned about laws, and so he writes down each and every regularity that his universe will display. In this case Saint Peter is left with the gargantuan task of arranging the initial properties in the universe in some way that will allow all God’s laws to be true together. The advantage to reductionism is that it makes Saint Peter’s job easier. God may nevertheless have chosen to be a metaphysical pluralist. (p. 33)

Cartwright’s position is an explicitly metaphysical one (see also Harré, 1993). A similar view of scientific matters of fact can be achieved through historical and sociocultural reflection (e.g., Dupré, 1993; Giere, 1999). A not-altogether unrelated debate in other realms of the philosophy of science is the extent to which theoretical constructs can be reified (e.g., van Fraassen, 1980; Hacking, 1983; Churchland & Hooker, 1985). This debate motivates reflection on which constructs are the real matters of fact of our mind/brain -- what cognitive structures and processes will be ascribed neurophysiological or some other genuine reality (in the history of psychology and linguistics, this has sometimes been termed the issue of "psychological reality"; e.g., Edelman, 2003; Fodor & Bever, 1965; Sapir, 1949).

Cartwright's reality toolbox is equally applicable in the case of the cognitive system. While many of these theses are intended to apply to science wholly, and often to physics specifically (e.g., Cartwright, 1999), cognitive scientific theories make a wonderful test case for this kind of philosophical position. It has become a hackneyed expression that the brain is "the most complex entity in the universe" -- and whatever that might mean, it certainly implies that theories of differing subject matter, admitting of distinctive nomological machines, are readily available in cognitive science.

The creation of nomological machines in psychological science is dependent entirely on the *units of analysis* selected for investigation. I offer a more detailed definition of unit of analysis below. But for now, a unit of analysis can be considered, as it is traditionally, a categorical or continuous scale of measurement of behavior -- from neuronal firing patterns, to Likert-scale ratings in a survey study. The unit of analysis is simply the set of attended behavioral events that satisfy the chosen system of measurement.

Consider a nomological machine using a unit of analysis whose measured behaviors are relatively complex (i.e., if considered in terms of underlying brain function): When a human participant rates the value of commodities when that participant has a status of owner or buyer (e.g., Loewenstein, 1996; Van Boven, Dunning, & Loewenstein, 2000). The experimental scenario generates regularities in relative value ratings as high or low depending on certain carefully established conditions. The unit of analysis here is judgment, and it comes to exhibit varying properties depending on complex socially relevant circumstances (i.e., owners rate the value of their property as greater than those seeking to purchase it). This social psychological nomological machine results in observations that give way to theories of "empathy gaps" that predict future patterns of social behavior in certain experimentally arranged settings. Moreover, as in Cartwright's wind-taken dollar bill scenario, it is in fact of long-standing concern whether or not such social behavioral regularities observed in laboratory settings are generalizable to broader real-world contexts (e.g., Mischel, 1968; Bem & Funder, 1978; and more recently Banaji & Crowder, 1989; Conway, 1991). This is unsurprisingly similar to the situation described above with physics and fluid dynamics: Are the fundamental regularities of social nomological machines applicable in contexts that violate their basic conditions -- such as in the complex sphere of real-world social behavior?

This same situation may be demonstrated across many units of analysis in psychological science. For example, the sometimes fascinating findings from neuro-imaging work (e.g., Haxby, Gobbini, Furey, Ishai, Schouten, & Pietrini, 2001) and whether they apply in broader theoretical and real-world contexts (e.g., Bub, 2000; Hassan, Nir, Levy, Fuhrmann, & Malach, 2004), or local sentence-processing effects (e.g., Frazier, 1987) and whether they hold in rich visual-world contexts (e.g., Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Nomological machines are thus continually refined and expanded. At each stage, they offer regularities that reveal the multifarious nature of the psychological universe -- yet are reliant upon the conveniently established conditions at any stage of refinement or expansion.

Cartwright's reality toolbox recommends that these models reveal the natures of the human cognitive system. The "nature" of the cognitive system is not inherently tied to one particular nomological machine. Our cognitive system instead reveals numerous different capacities, depending on the unit of analysis selected. All these are valid models of the functioning of our mind/brain.

Matter of fact #3: Spivey's echoes

It would be of little dispute to say that our behavior is guided by nervous system functioning, interacting with the surrounding stimulus energy. As Spivey and Dale (2004; see also Chapter 6) point out, both components of this closed system -- the interplay between brain-environment -- have strictly continuous characteristics. For example, the structure and function of neurons have a fractal structure, indicating self-similarity at any given temporal or spatial scale (Pellionisz, 1989; Teich, 1989). In addition, the nature of stimulus energies impinging on the cognitive system have a continuous character (Gibson, 1979).

Stopping at those observations offers only a simple reductionist solution to the matters of fact of cognition. Spivey and Dale (2004) go further in showing that continuity of this kind occurs even in higher-order cognitive processes, such as language. The topic of Chapters 2-5 was to reveal that there is some kind of continuity between categorization and response dynamics. This suggests that the matters of fact of the cognitive system involve some kind of continuous, dynamic representation and process. Such a process as categorization thus reveals "echoes" of continuity, simply by focusing on the dynamic properties in the task, such as response dynamics. As in the fractal patterns of neural structure and firing, fractal patterns in higher-order processes also reveal "echoes" in time-series analysis of reaction-time distributions (e.g., van Orden et al., 2001; Gilden, 2001). So the continuous-distributed framework is supported also by the character of higher-order cognitive processes.

A promising approach to showing that the matters of fact of cognition lie in continuity is to use the continuous interplay between processing and action as a litmus test, as exemplified in Chapters 2-5. This has been illustrated with categorization (Dale et al., in press), spoken-word recognition (Spivey et al., 2005), and sentence processing (Farmer, Cargill, Hindy, Dale, & Spivey, 2006). A straightforward problem with this approach is that it has not yet gathered extensive evidence -- but work in this direction has even begun to show echoes of continuity in the processing of logical and moral assessment of sentences (McKinstry, Dale, & Spivey, in preparation).

A second problem is deeper, and holds even if this reflection of continuity were evidenced through all cognitive processes. While continuity may hold between the process and its output in situations where competing responses are available, it may be that dynamical models that capture the proposed continuous process itself are unavailable. This presents a very important problem. A full account of a cognitive process would explain the variety of data on which it is based. For example, categorization behavior is guided by numerous constraints, such as prior knowledge or cultural domains (e.g., Burnett, Medin, Ross, & Blok, 2005; Medin, Ross, Atran, Cox, Coley, Proffitt, & Blok, 2005), or category-structure constraints (e.g., Murphy & Medin, 1985). A full account of this cognitive process would thus account also for its various inner-workings. While a distributed-continuous account may not be far off for categorization (e.g., the scaled-up system of Love et al., 2004), models for such processes as problem solving, moral reasoning, or social cognitive processes are not thoroughly attested (see, e.g., Balcetis & Dunning, 2005, for some ideas of how these models might proceed).

Nevertheless, the litmus test reveals an echo of continuous cognitive processing. Unfortunately, it cannot guarantee that continuous accounts for the process subjected to that litmus test will be comprehensive or even available. But if one accepts that the echoes may reveal underlying matters of fact, then it follows that there exists a sufficient and comprehensive dynamical account, somehow.

Summary reflections

None of these outlooks on cognitive matters of fact have full acceptance. Which do you espouse? Quinean confusion suggests that we should continue to strive for matters of fact, and push theory selection continuously as our theories come closer to reality (but perhaps never reach it). Cartwright's reality toolbox argues against the fundamentalist assumption that one theoretical framework is to be striven for: Each theory or model has its own unique contribution regarding matters of fact. Finally, Spivey's echoes serve as an enhanced kind of reductionism -- the low-level continuity of the nervous system along with evident continuity at higher levels suggests already that the continuous-distributed framework ought to be accepted as the cognitive matter of fact.

Quinean confusion and Cartwright's toolbox are not actually incompatible with Spivey's echoes. One might argue that underdetermination does not preclude rejection of whole families of theories. In that case, continuous-distributed theorists may argue that the underdetermination now applies only to competing continuous models, with the whole family of discrete ones rejected as unsuccessful. We can say the same in the case of Cartwright's metaphysical pluralism: For example, continuous-distributed theorists may make the metatheoretical (but also empirical) prediction that those natures or capacities will come to exhibit continuous and dynamic properties, rather than discrete-symbolic ones. Nevertheless, some argue that even ancient systems rejected long ago could be revamped to fit modern data, and appear more respectable to modern eyes, such as non-Copernican schemes for the solar system (Paul Thompson, personal communication, University of Toronto, 2000; see also Quine & Ullian, 1978; for a relevant yet far more interesting illustration of such an epistemological process, see Rokeach, 1964). Perhaps a compelling example from cognitive science is the perennial theoretical adjustment to generative grammar (Chomsky, 1965, 1986, 2000). In this theoretical evolution, not only are the details of the system reworked so as to preserve core theoretical principles while capturing novel data -- the framework has come to occasionally "shield" itself from other psychological data by concocting convenient dichotomies or conceptual schemes (e.g., as described above with I- and E-language).

Another important point is that Cartwright's perspective on models does not result in outright rejection of older ones that were successful in now "obsolete" nomological machines: "But that does not stop you from admitting that a crowbar is rigid, and, being rigid, is rightly described by Newton's laws; or that the solar system is composed of a small number of compact masses, and, being so composed, it too is subjected to Newton's laws." (p. 48)

Each recommendation for matters of fact is fairly radical in its broad assumption that the goal of science is to uncover the nature(s) of the universe, and that scientific success is measured by access to that nature or natures. Consider a fourth approach to cognitive matters of fact, one that at first pass might seem naive and obscurantist, but can be formulated into a thorough and perhaps enlightening perspective: Matters of fact don't matter. The next section limns the surface of such a metatheoretical approach.

Units of Analysis: Integrative Pluralism

John Dewey is described by Hilary Putnam (2006) as a prime mover of a potential pragmatic "third enlightenment." The first, says Putnam, was the Socratic and Platonic questioning of the nature of ethics and knowledge and its basis in Hellenic deities: "How do you know whether X is good because Zeus urges it, or that Zeus urges X because it is good?" The second, called the Enlightenment commonly, is the 17th century movement for scientific, ethical, and social "Reason." The third is what Putnam describes as a "critique of criticisms." The third enlightenment urges an endless stepping back from the systems and solutions to *all problems* (a set whose members include both "factual" and "ethical" problems), and "intelligently" critiquing them. As a "meta-critical" framework, this approach recommends a constant questioning of our systems of knowledge, in the hope of continuously modifying and advancing them.

In the realm of scientific knowledge, Dewey and Bentley (1948) lay out a broad epistemological framework for identifying observable variables, rendering them with labels or names, and specifying the interrelationships among them in scientific explanation. As we choose these labels and concoct explanations, paired with this *hope for epistemological progress*, they argue for some strong desiderata:

> The status of observation and the use of reports upon it are to be tentative, postulational, hypothetical. This condition excludes all purported materials and all alleged fixed principles that are offered as providing original and necessary "foundations"...The aim of observation and naming adopted is to promote further observation and naming which in turn will advance and improve. This condition excludes all namings that are asserted to give, or that claim to be, finished reports on "reality." (p. 113)

This perspective on scientific knowledge is directly related to cognitive explanation. In fact, Dewey and Bentley spend considerable time mulling over the observables, names, and systems of psychology. Their overall epistemology holds that a *transactional* understanding of systems is the fullest kind, and what science strives for. Transaction is distinguished from two other forms of understanding exemplified in the history of western scientific thinking: *self-action* and *inter-action*. The former seeks understanding of how our observables are self-guided -- that worldly entities from rocks to rock stars are self-moving, guided by their own independent underlying rules. The latter sees these entities as still independent, but importantly interacting in a system. Newtonian mechanics is such an example of how entities and their interrelationships can be explained. Transaction takes this further. The entities and their relationships cannot be considered real, separate things unto themselves, but are rather inherently intertwined in *our understanding*. This includes the entailment that the scientist, the knower herself, is part of this system: "One can easily 'think of' a world without a knower, or perhaps even of a knower without a world to belong to, and to know...[but a] 'real world' that has no knower to know it, has, so far as human inquiry is concerned...just about the same 'reality' that has the palace that in Xanadu Kubla Khan decreed.' (p. 142)

A defense of this epistemological system will not be made here. Instead, it can be argued that an application of this self-reflective system can cast a critique on how we talk about cognitive processes with a foundationalist's fervor. Dewey's and Putnam's (2006) pragmatic approach to science and ethics frames these epistemological systems in terms of *problems of knowing*, and they forcefully argue that progress in these problems is possible. This can be true without even beginning to concern ourselves with the supposed overall character of cognition -- which, it could be argued, is not ever part of the problems anyway.

For one, cognition is not a "natural kind." What we refer to as "cognition" is a complex assemblage of neuro-behavioral units of analysis, and the conditions under which they are studied. There are many ways of specifying units of measure for the behavior I am engaged in right now as I type this. For example, you might attach an electromyogram to my arm and digits and continuously track my movements from key to key. Instead, you might install some software on this computer to capture each explicit keystroke -- including the sequence of backspaces (there have been many). You might record the final product of these keystrokes -- this electronic document is such a data set. Through this final product, you might characterize the sentence structures by analysis, or the character of entire paragraphs. The units thereby become larger, in both space and time.

You could also describe the conditions in which this behavior occurs at varying levels of space and time. You could track the previous keystroke as a condition for the next, a relatively transient condition. You could measure my caffeine intake, my level of hunger, or other physiologically-relevant persistent conditions. Even more persistent conditions are my history and skill with keystrokes, or even further, my parents' proclivity for keystroking and whether there is any relevant heritability in their motor contributions to me. Such conditions can be considered for each unit of measure we choose. Complexity in space and time in the units we measure is compounded by the range of conditions we specify and relate to those units. And here we are just discussing keystroking. I engage in all kinds of behavior you might measure, in which this range of units and conditions holds.

The study of cognition starts here. Both the history, and the future, of cognitive science frames the selection of units and conditions. For example, Ebbinghaus' memory research supplied a vast and still-evolving framework for memory research. Identifying cognition, as a whole, with universal characterizations, "natural kind" descriptions, such as computation or information-processing (e.g., Dietrich, 1990; Fodor, 1993; Pylyshyn, 1984) is not where the questions and concerns of cognitive science lift off, nor necessarily get solved. For example, Fodor's (1983) characteristically cloudy conceptions of "mental states" are relevant only to a small range of phenomena. He offers intricately structured arguments out of fairly unusable notions of mentality -- a beautiful Ferrari made of paper: great on paper, but not on the road.

The branches of psychology, or cognitive science, pursue problems relevant to differing assortments of units and conditions. These problems are begot as much by a branch's history as by principled assessment of the composition of cognition. Each subfield of psychology has its own unique intellectual history. Any subfield, such as psycholinguistics or memory, is historically “bound,” or “enmeshed,” in an array of goals and questions regarding certain units of analysis and conditions. As mentioned above, Ebbinghaus shaped an early science of memory, framing the problems of memory in terms of certain behavioral units of measure (“recall” and “recognition”) in the context of particular experimental conditions (memorization stages and the nature of their stimulus items). Subsequent refinement of this problem seeks new ways of organizing the sources of recall and recognition (specific biographical events -- episodic memory; or errors of memory in thematic recall -- false memory research). Research questions about memory, enmeshed still in the behavioral units identified by Ebbinghaus, have been extended into real-world investigation of memory behavior -- ecological memory (e.g., Neisser & Winograd, 1995). As discussed further below, modern brain technologies permit these novel labels and proposed processes to be identified with particular regions of our brain. Each subsequent modification to the problems, still in some ways very much shaped by Ebbinghaus’ original formulations, has this historical fingerprint.

Similarly, the science of linguistics serves as an almost perfect example of this historical underpinning, with step-wise modification of questions and problems. Chomsky provided strong normative remarks on the goals of linguistic science. This normativity ensnares methods and any relevant evidence got from them. For example, Chomsky urged and still does an understanding of language acquisition from the perspective of an idealized and full-fledged adult speaker-hearer, largely on a basis of grammatical intuitions. The science of linguistics continues in this vein, while questions and problems of psycholinguistics have framed them differently -- questions of process and not just structure have led to a new history, and a new set of problems.

One could assert from these observations that matters of fact or "natures" play second-flute in the actual practice of psychological scientists. Solutions to problems better describe how psychological science progresses. Given the flexibility of behavioral unit selection, the sheer complexity of the human brain, and the vastly important role for a complex environment, solutions to these problems come in a variety of forms -- we should expect a plurality of explanatory schemes, rather than adopting a fundamentalist assumption regarding cognitive matters of fact.

However, within any given problem set (based on particular units of analysis, e.g., reaction time in the problem of sentence processing), there will inevitably be substantive debate. But this debate unfolds in these problem spaces. The very large and multifarious set of problems psychology addresses cannot be recast wholesale in terms of one or another such scheme. Yet these problem spaces do overlap. In other words, diverse psychological units of analysis and their conditions define problems that bear important relationships. This has invited extensive debate within cognitive science, but not enough integration. Chapter 6 has she some light on how seemingly disparate descriptive or explanatory schemes might be integrated through novel conceptual or technical advances.

What problem set was considered in Chapters 2-5? The unit of analysis selected was continuously measured motor output, and the problem was the time course of one possible process (the categorization process) and its relationship to this response dynamics. In this particular sphere, it seems that the continuous perspective offers the correct solution to this problem. That is, when we have selected a continuous temporal measure, the best characterization of the resultant behavior is one that sees a continuous flow of influence from cognition into action. But this does not solve any problems about how categories are formed, or how they are mulled over before issuing action. Chapters 2 to 5 have discussed some examples of theories that seek a solution to this, different, problem. The theory best accounting for this may be one that bears a strong theoretical relationship to the kind of continuity urged for here. But it might not -- and after a period of time in which another theory, of a seemingly distinct quality, holds reign, we may wish to seek another kind of dialectic to integrate understanding within broader domains. Pugilism can become unproductive. Instead, this dialectic would be based instead on this "pragmatic enlightenment": We can stand back, and with such tools as those presented in Chapter 6, consider ways in which continuous and non-continuous theories can relate.

Nevertheless, when we pay close attention to the time course of behavior, it seems that the eyes and the limbs show concomitant variation much like the underlying processes that guide them. With this unit of analysis, in this particular problem space, "cognition" and "action" seem intricately intertwined.

REFERENCES

Abrams, R., & Balota, D. (1991). Mental chronometry: Beyond reaction time. *Psychological Science*, *2*, 153-157.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419-439.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191-238.

Anderson, J. R. (1980). *Cognitive psychology and its implications*. New York, NY: W.H. Freeman.

Anderson, J. R. (1993). *Rules of the mind*. Erlbaum Associates.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washginton, DC: Winston.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications for a neural model. *Psychological Review*, *84*, 413-451.

Andrews, M. (2003). Unpublished Ph.D. dissertation, Cornell University.

Appelbaum, I. (1998). *Modularity*. In W. Bechtel and G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 625-635). Blackwell Publishers.

Ashby, F.G., Boynton, G., & Lee, W.W. (1994). Categorization response time with multidimensional stimuli. *Perception and Psychophysics*, *55*, 11-27.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K.W.Spence and J. T. Spence (Eds.), *The*

Psychology of Learning and Motivation: Advances in Research and Theory (Vol. 2, pp. 89-195). New York: Academic Press.

Atmanspacher, H. (2000). Ontic and epistemic descriptions of chaotic systems. In D. M. Dubois (Ed.), *Computing Anticipatory Systems: CASYS 99* (pp. 465-478). Berlin: Springer.

Balcetis, E., & Dunning, D. (2005). Judging for two: Some connectionist proposals for how the self informs and constrains social judgment. Invited chapter for M. Alicke, D. Dunning, & J. Krueger (Eds.), *Self and social judgment*. New York: Psychology Press.

Balkenius, C., & Grdenfors, P. (1991). Non-monotonic inferences in neural networks. In Allen J. A., Fikes R., and Sandewall, E. (Eds.), *Principles of knowledge representation and reasoning: Proceedings of the second international conference* (pp. 32-39). San Mateo, CA: Morgan Kaufman.

Ballard, D.H., Hayhoe, M.M., Pook, P.K., & Rao, R.P.N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723-742.

Balota, D., & Abrams, R. (1995). Mental chronometry: Beyond onset latencies in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1289-1302.

Balota, D., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language*, *28*, 14-36

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychology*, *44*, 1185-1193.

Barendregt, H. (1997). The impact of the lambda calculus in symbolic logic and computer science. *Bulletin of Symbolic Logic*, *3*, 181-215.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-660.

Bartlett, F.C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Bechtel, W. (1988). *Philosophy of science: An overview for cognitive science*. Hillsdale, NJ: Lawrence Erlbaum.

Bechtel, W. (1998). Representations and cognitive explanations: Assessing the dynamicist challenge in cognitive science. *Cognitive Science*, *22*, 295-318.

Bechtel, W., Abrahamsen, A., & Graham, G. (1998). *The life of cognitive science*. In W. Bechtel and G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 1-104). Blackwell Publishers.

Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: linking cognitive and neural states. *Philosophy of Science*, *66*, 175-207.

Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Eye Movements, Vol. 8 in Vision and visual Dysfunction*. CRC Press.

beim Graben, P. & Atmanspacher, H. (2004). Complementarity in classical dynamical systems. *Chaotic Dynamics*, arXiv: nlin.CD/0407046.

beim Graben, P. (2004). Incompatible implementations of physical symbol systems. *Mind and Matter*, *2*, 29-51.

beim Graben, P., Jurish, B., Saddy, D. & Frisch, S. (2004). Language processing by dynamical systems. *International Journal of Bifurcation and Chaos*, *14*, 599-621.

Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time: Assessing the personality of situation. *Psychological Review*, *85*, 485-501.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

Bitterman, M. E. (1960). Toward a comparative psychology of learning. *American Psychologist*, *15*, 704-712.

Block, N. (1980). Introduction: what is functionalism? In N. Block (Ed.), *Readings in philosophy of psychology* (pp. 171-184). Cambridge, MA: Harvard University Press.

Bloom, P. & Markson, L. (1998). Intentionality and analogy in children's naming of pictorial representations. *Psychological Science*, *9*, 200-204.

Bochenski, I. M. (1951). *Ancient formal logic*. North Holland Publishing.

Bollt, E., Stanford, T., Lai, Y., & Zyczkowski, K. (2000). Validity of threshold crossing analysis of symbolic dynamics from chaotic time series. *Physical Review Letters*, *85*, 3524-3527.

Bollt, E., Stanford, T., Lai, Y., & Zyczkowski, K. (2001). What symbolic dynamics do we get with A misplaced partition? On the validity of threshold crossings analysis of chaotic time-series. *Physica D*, *154*, 259-286.

Braitenberg, V. (1984). *Vehicles: experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Britten, K., Shadlen, M., Newsome, W., & Movshon, J. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*, 4745-4767.

Bub, D. N. (2000). Methodological issues confronting PET and fMRI studies of cognitive function: With special reference to Human Brain Function (1997). *Cognitive Neuropsychology*, *17*, 467-484.

Burnett, R.C., Medin, D.L., Ross, N.O., & Blok, S.V. (2005). Ideal is typical. *Canadian Journal of Experimental Psychology*, *59*, 3-10.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Cartwright, N. (1999). *The dappled world: a study of the boundaries of science*. Cambridge University Press.

Chater, N., Christiansen, M. H., & Reali, F. (2004). Is coevolution of language and language genes possible? Paper presented at the *Fifth International Conference on Language Evolution*. Leipzig, Germany.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1966). *Cartesian linguistics: a chapter in the history of rationalist thought*. New York, NY: Harper & Row.

chomsky 1968 (contra quine)

Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York, NY: Praeger Paperback.

Chomsky, N. (2000). *New horizons in the study of language and mind.* Cambridge University Press.

Christiansen, M.H. & Chater, N. (2001). Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences*, *5*, 82-88.

Churchland, P. S. (1986), *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge, MA: MIT Press.

Churchland, P. (1992). *A neurocomputational perspective: the nature of mind and the structure of science*. Cambridge, MA: MIT Press.

Clark, A. (1990). Connectionism, competence, and explanation. In M.A. Boden (Ed.), *The philosophy of artificial intelligence* (pp. 281-308). Oxford University Press.

Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. Oxford, UK: Oxford University Press.

Clark, H. (1993). *Arenas of language use*. Chicago, IL: University of Chicago Press.

Clark, K. B., Naritoku, D. K., Smith, D. C., Browning, R. A., & Jensen, R. A. (1999). Enhanced recognition memory following vagus nerve stimulation in human subjects. *Nature Neuroscience*, *2*, 94-98.

Cleeremans, A., Servan-Schreiber, D., & McClelland, J. (1989). Finite state automata and simple recurrent networks. *Neural Computation*, *1*, 372-381.

Coles, M.G.H., Gratton, G., Bashore, T.R., Eriksen, C.W., & Donchin. E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 529-553.

Collins, A., & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240-247.

Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407-428.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review* , *108*, 204-256.

Conway, M. A. (1991). In defense of everyday memory. *American Psychologist*, *46*, 19-26.

Cooper, L., & Shepard, R. (1973). The time required to prepare for a rotated stimulus. *Memory and Cognition*, *1*, 246-250.

Crain, S., & Steedman, M. (1985). On not being led up the garden path. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural Language Parsing*. Cambridge, MA: Cambridge

Cree, G.S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*, 371-414.

Crutchfield, J. P. & Young, K. (1989). Inferring statistical complexity. *Physical Review Letters*, *63*, 105-108.

Crutchfield, J.P. & Packard, N. (1982). Symbolic dynamics of one-dimensional maps: Entropies, finite precision, and noise. *International Journal of Theoretical Physics*, *21*, 433-466.

Crutchfield, J.P. & Young, K. (1990). Computation at the onset of chaos. In W. Zurek (Ed.), *Entropy, Complexity, and Physics of Information* (Vol. 8, pp. 223-269). Reading, MA: Addison-Wesley.

Crutchfield, J.P. (1994). The calculi of emergence: Computation, dynamics, and induction. *Physica D*, *75*, 11-54.

Crutchfield, J.P. (1998). Dynamical embodiments of computation in cognitive processes. *Behavioral and Brain Sciences*, *21*, 635-637.

Dale, R., Kehoe, C., & Spivey, M. (in press). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition*.

Dale, R. & Spivey, M.J. (2005). From apples and oranges to symbolic dynamics: A framework for conciliating notions of cognitive representations. *Journal of Experimental and Theoretical Artificial Intelligence, 17,* 317-342.

Damper, R. I. & Harnad, S. R. (2000). Neural network models of categorical perception. *Perception & Psychophysics*, *62*, 843-67.

Danovitch, J.H., & Keil, F.C. (2004). Should you ask a fisherman or a biologist?: Developmental shifts in ways of clustering knowledge. *Child Development*, *5*, 918 - 931.

Davidchack, R., Lai, Y., Klebanoff, A., & Bollt, E. (2001). Towards complete detection of unstable periodic orbits in chaotic systems. *Physics Letters A*, *287*, 99-104.

Daw, C., Finney, C., & Tracy, E. (2003). A review of symbolic analysis of experimental data. *Review of Scientific Instruments*, *74*, 916-930.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Dell, G.S., Lawler, E.N., Harris, H.D. & Gordon, J.K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*, 125-145.

Dennett, D. C. (1988). Quining qualia. In A. Marcel and E. Bisiach (Eds.) *Consciousness in Modern Science*. Oxford University Press.

Dennett, D. C. (1995). *Darwin's dangerous idea*. New York, NY: Simon & Schuester.

Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. Basic Books.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.

Devaney, D. (2003). *An introduction to chaotic dynamical systems* (2nd edition). Boulder, CO: Westview Press.

Dewey, J., & Bentley, A.F. (1948). *Knowing and the known*. Boston: Beacon.

Dietrich, E., & Markman, A.B. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind and Language*, *18*, 95-119.

Doyle M, & Walker R (2001). Curved saccade trajectories: Voluntary and reflexive saccades curve away from irrelevant distractors. *Experimental Brain Research*, *139*, 333-344.

Dreyfus, H. (1972). *What computers can't do: a critique of artificial reason*. New York, NY: Harper & Row.

Dreyfus, H. (1992). *What computers still can't do: a critique of artificial reason*. Cambridge, MA: MIT Press.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433-458.

Edelman, S. (2003). Generative grammar with a human face? *Behavioral and Brain Sciences*, *26*, 675-676.

Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. New York: Chapman & Hall

Elman, J.L. (1990). Finding structure in time. Cognitive Science, 14, 179-211.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness*. Cambridge, MA: The MIT Press.

Engbert, R., Scheffczyk, C., Krampe, R.T., Kurths, J. & Kliegl, R. (1998) Symbolic dynamics of bimanual production of polyrhythms. In: H. Kantz, J. Kurths & G. Mayer-Kress (Eds.), *Nonlinear analysis of physiological data*. Berlin: Springer.

Estes, Z. (2003). Domain differences in the struture of artificial and natural categories. *Memory and Cognition*, *31*, 199-214.

Farmer, T., Cargill, S. E., Hindy, N., Dale, R., & Spivey, M. J. (2006). Streaming x, y Coordinates Imply Continuous Interaction During On-line Syntactic Processing. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*.

Feigenson, L. & Carey, S. (2003). Tracking individuals via object files: Evidence from infants' manual search. *Developmental Science*, *6*, 568-584.

Feldman, D.P. & Crutchfield, J.P. (1998). Measures of statistical complexity: Why? *Physics Letters A*, *238*, 244-252.

Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003-1015.

Fialkowski, K. R. (1986). A mechanism for the origin of the human brain: A hypothesis. *Current Anthropology*, *27*, 288-290.

Flanagan, J.R. & Lolley, S. (2001). The inertial anisotropy of the arm is accurately predicted during movement planning. *Journal of Neuroscience*, *21*, 1361-1369.

Fodor, J.A. (1975). *The language of thought*. New York, NY: Cromwell.

Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive science*. Cambridge, MA: MIT Press.

Fodor, J.A. (1983). *Modularity of mind*. Cambridge, MA: Bradford.

Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor, J. A. (1993). The persistence of the attitudes. In S. M. Christensen, & D. R. Turner (Eds.), *Folk psychology and the philosophy of mind* (pp. 221-246). Erlbaum.

Fodor, J.A. (2000). *The mind doesnt work that way*. Bradford Books.

Fodor, J. A., & Bever, T. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, *4*, 414-420.

Fodor, JA, & Pylyshyn, ZW (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3-71.

Frazier, L. (1987). Theories of sentence processing. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 291-307). MIT Press.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, *24*, 437-468.

French, R. K. (1978). The thorax in history 1. From ancient times to Aristotle. *Thorax*, *33*, 10-18.

Garrod, S. & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, *16*, 77-90.

Gauker, C. (2003). *Words without meaning*. Cambridge, MA: MIT Press.

Gaveau, V., Martin, O., Prablanc, C., Pelisson, D., Urquizar, C., & Desmurget, M. (2003). On-line modification of saccadic eye movements by retinal signals. *Neuroreport*, *14*, 875-878.

Georgopoulos, A. (1995). Motor cortex and cognitive processing. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 507-517). Cambridge, MA: The MIT Press.

Georgopoulos, A., Lurito, J., Petrides, M., Schwartz, A., & Massey, J. (1989). Mental rotation of the neuronal population vector. *Science*, *243*, 234-236.

Gibson, J. J. (1979). *The ecological approach to perception*. Boston: Houghton.

Gilden, D.L. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, *108*, 33-56.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558-565.

Glenberg, A.M., & Robertson, D.A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*, 379-401.

Goertzel, B. (1998). Learning the language of mind: Symbolic dynamics for modeling adaptive behavior. In C. Wynne, & J. Staddon (Eds.), *Models of action: Mechanisms for adaptive behavior* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Gold, J., & Shadlen, M. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, *404*, 390-394.

Gold, J., & Shadlen, M. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*, 10-16.

Goldberger A.L., Amaral L.A.N., Glass L., Hausdorff J.M., Ivanov P.Ch., Mark R.G., Mietus J.E., Moody G.B., Peng C.K., Stanley H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, *101*, e215-e220.

Goodale, M.A., Pelisson, D., & Prablanc, C. (1986). Large adjustments in visual guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, *320*, 748-750.

Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge, MA: Harvard University Press.

Green, C.D. (1998). Are connectionist models theories of cognition?*Psycoloquy*, *9* (4).

Green, C.D. (2001). Scientific models, connectionist networks, and cognitive science. *Theory & Psychology*, *11*, 97-117.

Hadley, R. (1994) Systematicity revisited. *Mind and Language*, *9*, 431-444.

Harm, M.W. & Seidenberg, M.S. (1999) Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491-528

Harnad, S. (Ed.) (1987). *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Hassan U, Nir Y, Levy I, Fuhrmann G, & Malach R (2004) Intersubject. synchronization of cortical activity during natural vision. *Science*, *303*, 1634-1640

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425-2430.

Hayes, S.C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: a post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic.

Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1264-1282.

Hinton, G.E. & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, *98*, 74-95.

Hirahara, M. & Nagano, T. (2003). A neural network for the typicality effects. *Proceedings of the International Joint Conference on Neural Networks*, 4, 2502-2505.

Horne, P.J. & Lowe, C.F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 185-241.

Hume, D. (1748/1967). *An enquiry concerning the human understanding*. In E. A. Burtt (Ed.), *The English Philosophers from Bacon to Mill* (pp. 585-689). New York, NY: Random House, Inc.

Job, R., Rumiati, R., & Lotto, L. (1992). The picture superiority effect in categorization: visual or semantic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1019-1028.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, *13*, 289-303.

Edman, I. (1928). *The works of Plato*. Modern Library Books.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, *39*, 170-210.

Keller, H. (1903). *Story of my life*. W. W. Norton & Company.

Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kennel, M.B. & Buhl, M. (2003). Estimating good discrete partitions from observed data: symbolic false nearest neighbors. *Physical Review Letters*, *91*, 084102.

Kneale, W. (1948). Boole and the revival of logic. *Mind*, *57*, 149-175.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2004). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*, 95-127.

Kuhn, T. (1963). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Kurths, J., Voss, A., Saparin, P., Witt, A., Kleiner, H. & Wessel, N. (1995). Quantitative analysis of heart rate variability. *Chaos*, *5*, 88-94.

Lachter, J. & Bever, T. (1988). The relation between linguistic structure and associative theories of language learning : A constructive critique of some connectionist learning models. *Cognition*, *28*, 195-247.

Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge University Press.

Lake, D.E., Richman, J.S., Griffin, M.P., & Moorman, J.R. (2002). Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology*, *283*, R789-R797.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* Basic Books.

Lamberts, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, *124*, 161-180.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227-260.

Lamberts, K. (2002). Feature sampling in categorization and recognition of objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *55A*, 141-154.

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211-240.

Lasnik, H. (2002). The minimalist program in syntax. *Trends in Cognitive Sciences*, *6*, 432-437.

Leahey, T.H. (2001). *A history of modern psychology*. Upper Saddle River, NJ: Prentice-Hall.

Liberman, A. (1982). On finding that speech is special. *American Psychologist*, *37*, 148-167.

Liberman, A., Harris, K., Hoffman, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *53*, 358-368.

Love, B., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*, 309-332.

Lukasiewicz, J. (1951). *Aristotle's syllogistic*. Oxford: Clarendon Press.

Lynch, E.B., Coley, J.D., & Medin, D.L. (2000). Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices. *Memory and Cognition*, *28*, 41-50.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.

Magnuson, J. S. (2005). Moving hand reveals dynamics of thought: Commentary on Spivey, Grosjean, & Knoblich (2005). *Proceedings of the National Academy of Sciences*, *102*, 9995-9996.

Magnuson, J., Tanenhaus, M., Aslin, R., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*, 202-227.

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*, 97-115.

Markman, A.B., & Dietrich, E. (2000). In defense of representation. *Cognitive Psychology*, *40*, 138-171.

Marr, D. (1982). *Vision.* New York, NY: Henry Holt & Co.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71-102.

Massaro, D.W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, *21*, 398-421.

Massaro, D.W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. MIT Press/Bradford.

Mattes, S., Ulrich, R., & Miller, J. (2002). Response force in RT tasks: Isolating effects of stimulus probability and response probability. *Visual Cognition*, *9*, 477-501.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1-86.

McClelland, J., L. & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310-322.

McMurray, B., Tanenhaus, M., & Aslin, R. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*, B33-B42.

McMurray, B., Tanenhaus, M., Aslin, R., & Spivey, M. (2003). Probabilistic constraint satisfaction at the lexicałphonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research*, *32*, 77-97.

McMurray, R., & Spivey, M. (1999). The categorical perception of consonants: The interaction of learning and processing. In *Proceedings of the Chicago Linguistic Society*.

McRae, K., & Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283-312.

McRae, K., de Sa, V.R., & Seidenberg, M.S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99-130.

Medin, D. L. (1989). Concepts and concept structure. *American Psychologist*, *44*, 1469-1481.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, Vol. 5 (pp. 189-223). Hillsdale, NJ: Lawrence Erlbaum Associates.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207-238.

Medin, D. L., & Schaffer, M. M. (1984). Concepts and concept formation. *Annual Review of Psychology*, *35*, 113-138.

Medio, A., & Lines, M. (2001). *Nonlinear dynamics: A primer*. Cambridge University Press.

Meehl, P.E. (1998). The power of quantitative thinking. Address at the American Psychology Society, May. Available online: http://www.tc.umn.edu/~pemeehl

miller, 1998 (phil of lang)

Miller, J. & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology*, *46*, 101-151.

Miller, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 273-296.

Moore, E.F. (1956). Gedanken-experiments on sequential machines In C. E. Shannon and J. McCarthy (Eds.), *Automata Studies* (pp. 129-153). Princeton: Princeton University Press.

Morse, M. & Hedlund, G.A. (1938). Symbolic dynamics. *American Journal of Mathematics*, *60*, 815-866.

Murphy, G. L. (1991). Parts in object concepts: Experiments with artificial categories. *Memory and Cognition*, *19*, 423-438.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, G. L. & Medin, D.L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289-316.

Nakayama, K., & Joseph, J. (1998). Attention, pattern recognition, and pop-out visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279-298). Cambridge, MA: MIT Press.

Nederhouser, M. & Spivey, M.J. (2004). Simulating the temporal dynamics of categorization. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Lawrence Erlbaum.

Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Newell, A., & Simon, H. (1956). The logic theory machine: A complex information processing system. *IEEE Transactions on Information Theory, 2*, 61-79.

Newell, A., & Simon, H. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*, 113-126.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151-166.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 87-108.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*, 700-708.

Nosofsky, R. M. (1991). Typicality in logically defined categories: Exemplar-similarity versus rule instantiation. *Memory and Cognition, 19*, 131-150.

Nosofsky, R. M., & Palmeri, T.J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266-300.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 924-940.

Olds, E., Cowan, W., & Jolicoeur, P. (2000). The time-course of pop-out search. *Vision Research*, *40*, 891-912.

Oliphant, M. (2002). Rethinking the language bottleneck: Why don't animals learn to communicate? In K. Dautenhahn & C.L. Nehaniv (Eds.), *Imitation in animals and artifacts. Complex adaptive systems* (pp. 311-325). Cambridge, MA: The MIT Press.

Osman, A., Kornblum, S., & Meyer, D. (1986). The point of no return in choice reaction time: Controlled and ballistic stages of response preparation. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 243-258.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255-287.

Palmer, D. (2003). Specifying psychology's observable units: Toward an integration of Kantor's behavior segment, Skinner's operant, and Lee's deed. *Behavior and Philosophy*, *31*, 81-110.

Pellionisz, A.J. (1989). Neural geometry: Towards a fractal model of neurons. In R.M. Cotterill (Ed.), *Models of Brain Function* (pp. 453-464). Cambridge University Press.

Perrett, D., Oram, M., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalization of recognition without mental transformations. *Cognition*, *67*, 111-145.

Pinker, S. (1997). How the mind works. New York, NY: Norton.

Pisoni, D., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics*, *15*, 285-290.

Pitt, M. A., Myung, I., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*, 472-491.

Plaut, D.C. & Gonnerman, L. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*, 445-485.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*, 603-619.

Pollack, J. B. (1991). The induction of dynamical recognizers. *Machine Learning*, *7*, 227-252.

Port, R. & van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

Posner, M. L. (1980). The orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.

Posner, M. L., & Petersen, S. E. (1990). The attenion system of the human brain. *Annual Review of Neuroscience*, *13*, 25-42.

Pulvermueller, F. (1999). Word's in the brain's language. *Behavioral and Brain Sciences*, *22*, 253-336.

Putnam, H. (1960). *Minds and machines*. In S. Hook (Ed.), *Dimensions of mind* (pp. 138-164). New York: Collier.

Putnam, H. (1967). Psychological predicates. In W. H. Capitan and D. D. Merrill (Eds.) *Art, mind, and religion* (pp. 37-48). Pittsburgh, University of Pittsburgh Press.

Putnam, H. (2006). *Ethics without ontology*. Harvard University Press.

Pylyshyn, Z. (2000). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*, 341-365.

Quine, W.V. & Ullian, J.S. (1970). *The web of belief*. New York, NY: Random House.

Quine, W.V. (1960). *Word and object*. Cambridge, MA: MIT Press.

quine, 1970

Rayside, D., & Kontogiannis, K. (2001). On the Syllogistic Structure of Object-Oriented Programming. In the proceedings of *The 23rd International Conference on Software Engineering*.

Redington, M. & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General*, *125*, 123-138.

Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382-407.

Rehder, B., & Hoffman, A. B. (2005). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 811-829.

Rehder, B., & Murphy, G. L. (2003). A knowledge-resonance (KRES) theory of category learning. *Psychonomic Bulletin and Review*, *10*, 759-784.

Reynolds, J., & Desimone, R. (2001). Neural mechanisms of attentional selection. In J. Braun, C. Koch, & J. Davis (Eds.), *Visual attention and cortical circuits* (pp. 121-135). Cambridge, MA: The MIT Press.

Richardson, D., Dale, R., & Spivey, M. (in press). Eye movements in language and cognition: A brief introduction. In Gonzalez-Marquez, Coulson, Mittelberg, & Spivey (Eds.) *Methods in Cognitive Linguistics*. John Benjamins.

Richman J.S. & Moorman J.R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology: Heart and Circulatory Physiology*, *278*, H2039-2049.

Rips, L., Shoben, E.J., & Smith, E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1-20.

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*, 31-40.

Roberts, S. & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*, 358-367.

Rodriguez, P., Wiles, J., & Elman, J. (1999). A recurrent neural network that learns to count. *Connection Science: Journal of Neural Computing, Artificial Intelligence and Cognitive Research*, *11*, 5-40.

Rohde, D. (2002). *A Connectionist Model of Sentence Comprehension and Production*. Unpublished Ph.D. thesis, Carnegie Mellon University.

Rolls, E., & Tovee, M. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, *73*, 713-726.

Rosch, E. & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.

Rosch, E. (1973). Natural categories. *Cognitive Psychology*, *4*, 328-350.

Rosch, E. (1975) Cognitive representations of semantic categories. *Journal of Experimental Psychology*, *104*, 192-233.

Rosen, D. (1975). An argument for the logical notion of a memory trace. *Philosophy of Science*, *42*, 1-10.

Rosenbaum, D. (2005). The Cinderella of psychology: The neglect of motor conrol in the science of mental life and behavior. *American Psychologist*, *60*, 308-317.

Rueckl, J.G. (2002). The dynamics of visual word recognition. *Ecological Psychology*, *14*, 5-19.

Rumelhart, D.E. & McClelland, J.L. (1986). *Parallel distributed processing: Explorations in the microstrcuture of cognition*. Cambridge, MA: MIT Press.

Schall, J. (2000). Decision making: From sensory evidence to a motor command. *Current Biology*, *10*, R404-R406.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: an inquiry into human knowledge structures*. Erlbaum.

Sheliga, B., Craighero, L., Riggio, L. & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research*, *114*, 339-351.

Sheliga, B., Riggio, L., & Rizzolatti, G. (1995). Spatial attention and eye movements. *Experimental Brain Research*, *105*, 261-275.

Shin, J.C. & Rosenbaum, D.A. (2002). Reaching while calculating: Scheduling of cognitive and perceptual-motor processes. *Journal of Experimental Psychology: General*, *131*, 206-219.

Skinner, B.F. (1938). *The behavior of organisms: an experimental analysis*. New York, NY: Appleton-Century-Crofts.

Smith, J. C. (1991). *Historical foundations of cognitive science*. Springer.

Smith, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science*, *13*, 437-442.

Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 3-27.

Smolensky, P. (2001). Grammar-based connectionist approaches to language. In M. H. Christiansen & N. Chater (Eds.), *Connectionist psycholinguistics* (pp. 319-347). Westport, CT: Ablex.

Snodgrass, J.G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior*, *23*, 3-22.

Snodgrass, J.G. & McCullough, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 147-154.

Spivey-Knowlton, M.(1996). *Integration of visual and linguistic information: Humandata and model simulations*. Ph.D. Thesis, University of Rochester, 1996.

Tanenhaus, M. K., Spivey-Knowlton, M. J., & Hanna, J. E. (2000). Modeling thematic and discourse context effects with a multiple constraints approach: implications for the architecture of the language comprehension system. In M. W. Crocker, M. J. Pickering, & C. Clifton

(Eds.), *Architectures and mechanism for language processing* (pp. 90–118). Cambridge UniversityPress.

Spivey, M. J. (2006). *The continuity of mind*. Oxford University Press.

Spivey, M.J., & Dale, R. (2004). The continuity of mind: Toward a dynamical account of cognition. In B. Ross (Ed.), *Psychology of Learning and Motivation* v.45 (pp. 85-142). Elsevier Academic Press.

Spivey, M., J. & Marian, V. (1999). Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*, 281-284.

Spivey, M., J. & Tanenhaus, M. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 1521-1543.

Spivey, M., J. Grosjean, M., Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*, 10393-10398.

Spivey, M. J., Tanenhaus, M., Eberhard, K., & Sedivy, J. (2002b). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447-481.

Spivey, M. J., Tyler, M., Eberhard, K., & Tanenhaus, M. (2001). Linguistically mediated visual search. *Psychological Science, 12*, 282-286.

Spivey, M.J. & Dale, R. (2004). The continuity of mind: Toward a dynamical account of cognition. In B. Ross (Ed.), *Psychology of learning and motivation*, Vol. 45 (pp. 85-142). Amdsterdam: Elsevier.

Spivey-Knowlton, M. J., & Saffran, J. (1995). Inducing a grammar without an explicit teacher: Incremental distributed prediction feedback. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Spivey-Knowlton, M. J., Sedivy, J., Eberhard, K. & Tanenhaus, M. (1994). Psycholinguistic study of the interaction between language and vision. In *Proceedings of the 12th National Conference on Artificial Intelligence: Workshop on the Integration of Natural Language and Vision Processing* (pp. 189-192). Menlo Park, CA: AAAI Press.

Steedman, M. (2001). Connectionist sentence processing in perspective. In M. H. Christiansen & N. Chater (Eds.), *Connectionist psycholinguistics* (pp. 348-372). Westport, CT: Ablex.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, *23*, 295-322.

Stokes, P. (2002). *Philosophy: 100 essential thinkers.* Arcturus Publishing Limited.

Storms, G. (2005). Exemplar models in the study of natural language concepts. In B. Ross (Ed.), *Psychology of Learning and Motivation*, Vol. 45 (pp. 1-39). Amsterdam: Elsevier.

Sun, R. (1997). Learning, action and consciousness: A hybrid approach toward modeling consciousness. *Neural Networks*, *10*, 1317-1331.

Tabor, W. (2000). Fractal encoding of context-free grammars in connectionist networks. *Expert Systems: The International Journal of Knowledge Engineering and Neural Networks*, *17*, 41-56.

Tabor, W. (2002). The value of symbolic computation. *Ecological Psychology*, *14*, 21-51.

Tabor, W., & Tanenhaus, M. (1999). Dynamical models of sentence processing. *Cognitive Science*, *23*, 491-515.

Tabor. W. (2001). Dynamical assessment of symbolic processes with backprop nets. In *Proceedings of the INNS-IEEE International Joint Conference on Artificial Neural Networks*.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, *268*, 1632-1634.

Teich, M. C. (1989). Fractal character of the auditory neural spike train. *IEEE Transactions on Biomedical Engineering*, *36*, 150-160.

Tenenbaum, J.B. & Griffiths, T.L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, *24*, 629-664.

Thelen, E. & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Tipper, S., Howard, L., & Jackson, S. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition*, *4*, 1-38.

Tooby, J., & Cosmides, L.. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, and J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 19-136). New York, NY: Oxford University Press.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201-237.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

Tsuda, I. (2001). Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behavioral and Brain Sciences*, *25*, 793-848.

Tucker, M., & Ellis, R. (1998). On the relations of seen objects and components of potential actions..*Journal of Experimental Psychology: Human Perception and Performance*, *24*, 830-846.

Ulrich, R., Mattes, S., & Miller, J. (1999). Donders's assumption of pure insertion: An evaluation on the basis of response dynamics. *Acta Psychologica*, *102*, 43-75.

Underwood, G. (2005) (Ed.), *Eye Guidance in Cognition*. Oxford: Oxford University Press.

Uttal, W. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: The MIT Press.

Van Fraassen, B.C. (1980). *The scientific image*. Oxford University Press.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*, 615-665.

Varela, F., Thompson, E.T., & Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Verbeemen, T., Storms, Gert, & Verguts, T. (2003). Determinants of speeded categorization in natural concepts. *Psychologica Belgica*, *43*, 139-151.

Viswanathan, M. & Childers, T.L. (2003). An enquiry into the processing of categorization of pictures and words. *Perceptual and Motor Skills*, *96*, 267-287.

Ward, L. (2002). *Dynamical cognitive science*. Cambridge, MA: MIT Press.

Weizenbaum, J. (1966). ELIZA A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*, 36-45.

Williams, S.G. (2002). Introduction to symbolic dynamics. In S.G. Williams (Ed.), *Proceedings of Symposia in Applied Mathematics: Symbolic Dynamics* *and its Applications* (Vol. 60, pp. 1-13). Providence, RI: American Mathematical Society.

Winograd, T. (1972). *Understanding natural language*. Oxford, UK: Academic Press.

Wolfe, J. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*, 33-39.

Young, R.M. & Lewis, R.L. (1999). The Soar cognitive architecture and human working memory. In P. Shah & A. Miyake (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 224-256). New York, NY: Cambridge University Press.