COGNITIVE PROCESSES: A WHITEHEADIAN PERSPECTIVE

Presented by Jacob Hoke Sansom

in partial fulfillment of the requirements for completion of the
Evidence and Inquiry certificate and the
Polymathic Scholars honors program in the College of Natural Sciences at
The University of Texas at Austin

Spring 2020

May 12, 2020

Michael Mauk, Ph.D.
Thesis Supervisor
Department of Neuroscience
The University of Texas at Austin

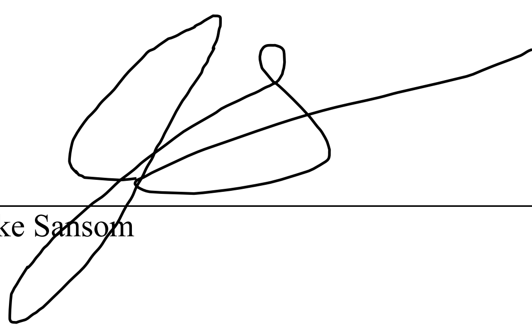
Date

Rebecca A. Wilcox, Ph.D.
Second Reader
CNS Honors
The University of Texas at Austin

Date

I intend to submit a copy of my Polymathic Scholars thesis to the Texas ScholarWorks (TSW) Repository. For more information on the TSW, please visit https://repositories.lib.utexas.edu/.

Cognitive Processes: A Whiteheadian Perspective



Jake Sansom

May 13, 2020

Date

Table of Contents

Acknowledgements

This thesis would not have been possible without the innumerable contributions, intellectual and otherwise, of those who guided me throughout the writing process. First and foremost, Dr. Michael Mauk consistently encouraged me to pursue my interests no matter how bizarre and reaching. Dr. Rebecca Wilcox, adopting a different but equally valuable approach, taught me to transform these ideas into something achievable. I am thankful for my family, who has consistently believed in me even when this process seemed quite bleak. I am especially grateful for my father, whose philosophical company first gave me the idea for this thesis. Finally, I would like to thank my friend Matthew Shuirman for acting as a sounding board for my ideas even when they reached well outside his scope of interest.

Abstract

Although interdisciplinary analyses bridging philosophy and science can be highly valuable, methodological differences between the two fields make such analyses rare. A prime example of this phenomenon is the study of cognition, which is the mental process giving rise to human thought. Philosophers of mind and cognitive scientists both work to elucidate the properties of cognition; however, they employ vastly disparate tools. This paper synthesizes process metaphysics—a particular branch of philosophy based on the thinking of Alfred North Whitehead and philosophers following in his tradition—and cognitive science to demonstrate the value of their combination and to suggest lucrative directions for future study. More specifically, the paper uses the perspectives and tools of Whiteheadian process metaphysics to critique *Categorizing Cognition* and “Design for a Working Memory,” two works that illuminate a particular approach under the symbolic paradigm of cognitive science. To do this, the paper begins by summarizing relevant cognitive science challenges, introducing difficulties inherent in answering the question “what is cognition?”, as well as the currently operative definitions and competing frameworks, and concludes with special emphasis on a thought ranking approach built by *Categorizing Cognition* and based on the working memory model of “Design for a Working Memory.” Next, the paper summarizes the relevant tenets of Whiteheadian process metaphysics and its notion of “actual entities.” Building upon this foundation, subsequent analysis of cognitive science and process metaphysics together identifies analogs between the two fields to confirm their mutual focus and build a foundation for cross-disciplinary discussion. The two analytical frameworks of process metaphysics—genetic and coordinate analysis—are used to critique the approach of the cognitive science texts. Ultimately, the paper demonstrates the firm opposition of process metaphysics to certain elements within symbolic models of cognition. By reevaluating the assumptions underlying these models, the paper identifies novel directions for future study.

Key Words: cognition, process, metaphysics, philosophy, Whitehead, symbol, classical, grounding problem

Introduction

By historical standards, neuroscience and cognitive science are both in their infancies. Neuroscience was not a cohesive field until the second half of the 1900's, and cognitive science evolved even later (Kandel & Squire, 2000; Miller, 2003). It is only natural for the novelty of these fields to produce multiple schools of thought stuck in mutual disagreement. How, then, can neuroscientists and cognitive scientists overcome this gridlock to make efficient progress?

In his book *The Structure of Scientific Revolutions*, the philosopher Thomas Kuhn surveys a variety of scientific achievements throughout history to answer this very question. Primarily, he argues that efficient scientific progress results when scientists adopt a particular group of assumptions. He coins the term "paradigm" to refer to such groups of assumptions. He comments, "In the absence of a paradigm or some candidate for paradigm, all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant . . . it produces a morass" (Kuhn, 1962, p. 15). In other words, science utilizes a hypothesis-driven scheme to increase its rate of progress. Philosophical and theoretical work leads to the creation of a paradigm within which findings can be efficiently generated.

In contrast, Alfred North Whitehead—one of the primary authors this essay will focus on—argues that this iterative, hypothesis and experiment cycle is largely unnecessary. He describes, "If my view of the function of philosophy is correct, it is the most effective of all the intellectual pursuits. It builds cathedrals before the workmen have moved a stone, and it destroys them before the elements have worn down their arches" (Whitehead, 1925, p. 8).

Although Whitehead's view on the scientific method was extreme, he drew conclusions similar to those of Thomas Kuhn. Primarily, both philosophers argued that the scientific method depends upon assumptions. The modification of these assumptions can yield novel directions for experimentation and discovery. Kuhn described the historical alteration of assumptions as an iterative process, instantiated when a scientific paradigm can no longer account for its own experimental results. Such a situation causes a “crisis” that requires scientists to reevaluate. Whitehead, in contrast, believed rigorous philosophical analysis could tease out errors within these assumptions, providing a starting place more quickly than the scientific method ever could.

With this purpose in mind, Whitehead worked to develop a watertight philosophical scheme. In 1929, he published *Process and Reality*, a philosophical scheme meticulously constructed from first principles. Whitehead’s express goals incentivized him to hold the arguments within *Process and Reality* to heightened standards of rigor.

Rather than fastening to the beliefs of one philosopher or the other, this paper takes a balanced position. Primarily, it aims to partially fulfill Whitehead’s original goal by using his philosophy to examine the assumptions of a particular scientific paradigm. The end goal, however, is not an abandonment of science, but instead the identification of new directions for study. Given its youth, cognitive science is ripe for philosophical evaluation. A scarcity of experimental evidence forces paradigms within cognitive science to fully leverage assumptions for progress. The reevaluation of these assumptions could be a fruitful undertaking.

More specifically, this paper aims to analyze the assumptions underlying symbolic theories of cognition—a particular paradigm within cognitive science—through the lens of Whiteheadian philosophy. As it is well beyond the scope of the paper to evaluate the paradigm as a whole, two texts were chosen for analysis. *Categorizing Cognition*, a (2014) book written by Drs. Graeme Halford, William Wilson, Glenda Andrews, and Steven Phillips, works well for the paper because its intent was to draw upon work from across the paradigm. Because *Categorizing*

Cognition synthesizes such a broad array of scientific literature, its critiques will be more likely to find broad application. Secondly, Dr. Klaus Oberauer's "Design for a Working Memory" is a focal point in *Categorizing Cognition*, so the paper references it frequently. Finally, the paper depends upon the summary work of Dr. Elizabeth Kraus in *The Metaphysics of Experience* to accurately portray Whitehead's philosophy.

The paper begins by introducing cognitive science as a field then describes the sub-region of this field *Categorizing Cognition* and "Design for a Working Memory" occupy. After doing so, the paper provides a detailed summary of Alfred North Whitehead's process metaphysics. The rest of the paper analyzes the two cognitive science texts jointly through the perspective of process metaphysics. Using what Whitehead terms the "genetic analysis" of actual entities, the paper argues that symbolic theorists have not yet solved the decades-old "symbol grounding problem." Similarly, using an alternative mode of analysis Whitehead terms "coordinate analysis," the paper argues that the solution to the symbol grounding problem—among other issues—may be an alternative long term memory architecture. Suggestions provided by the application of Whitehead's genetic and coordinate analyses should not be misinterpreted as concrete solutions. Rather, the paper aims to provide a fresh perspective that can hopefully serve as a foundation for novel research in the future.

What is Cognition?

Everyone has an intuitive sense as to what cognition is. In its most basic definition, it is the mental process giving rise to human intellectual activity. However, any simple definition of cognition underestimates its complexity. To fully understand the elusive concept that is "cognition," it is necessary to study the particular difficulties faced by experimental cognitive scientists.

By definition, the scientific method lends itself to an iterative development of theories to accurately explain observations. Oftentimes, therefore, theoretical accounts are limited by the scientist's ability to observe the phenomena they are studying. Gregor Mendel, for example, crafted an early model of genetic inheritance by successfully manipulating allele propagation during pea plant reproduction (Gayon, 2016). Due to his inability to observe and manipulate chromosomes, however, Mendel's theory failed to fully capture the intricacies of genetic inheritance (Gayon, 2016).

Despite advancements, scientific tools for the observation of cognitive processes are relatively primitive. *In vivo* studies of active cognitive processes in humans are impossible at a high resolution (single-neuron scale). Functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) data indicate general areas of high neural activation, but they do not allow scientists to observe the behavior of individual neurons (Kandel and Squire, 2000). For this reason, cognitive scientists must rely on intuition from tangentially related disciplines and the direct observation of low-resolution data. Such knowledge comes from a diversity of fields, not limited to psychology, neuroscience, linguistics, and computer science (Hardcastle, 1996; Miller, 2003).

Definitions

Jerry Fodor, a philosopher and cognitive scientist, formulated an early definition of cognition. Cognition, he argued, is comprised of the "sorts of systems that people have in mind when they talk, pretheoretically, of such mental processes as thought and problem-solving" (Fodor, 1983, as cited in Murphy, 2019). Although it was probably the strictest at the time, Fodor's definition left much room for growth. The meaning of "thought," for example, seems clear at first, but it is obscured by a scarcity of explicit characteristics. The vagueness of the definition provided no measurable goals by which a framework could be evaluated. Additionally, it is unclear whether thought and problem-solving are the only two phenomena required to produce sentient thought.

With good reason, therefore, scientists are constantly working to sharpen their definition of cognition. Over time, many other phenomena have been added to the list of hypothesized cognitive features. G. L. Murphy (2019) summarized a recent list as follows: "the term higher-level cognition has come to refer to such things, normally including processes such as categorization, decision making, planning, and discourse comprehension in addition to those Fodor mentioned" (p. 4).

Clearly, cognitive scientists have exerted great effort to lengthen and clarify the list of cognitive features. Still, however, these efforts have failed to sufficiently restrict the range of possible cognitive architectures. Paradigmatic fissures within cognitive science are therefore quick to form and difficult to resolve.

Competing Frameworks

The paradigmatic study of cognition is generally divided over methodologies as much as results. Within his 2010 paper "Probabilistic models of cognition: Exploring representations and inductive biases," Dr. Thomas Griffiths works with collaborators to partition cognitive science methodologies into two main categories: the "mechanism-first" strategy and the "function-first" strategy. The mechanism-first school studies the intricacies of neurons and neural networks (both biological and artificial) to create architectures that seem fundamentally cognitive. This methodology was popularized in the connectionist paradigm, named thus after its focus on the connections inherent in large neural network models. The function-first school, conversely, begins "with abstract principles that allow agents to solve problems posed by the world—the functions that minds perform—and then attempt[s] to reduce these principles to psychological and neural processes" (p. 357). In other words, the former expects cognition to "emerge" as neural network models become more lifelike, whereas the latter expects the two-step process of defining cognitive qualities and more accurately implementing them to be a more efficient process.

Although much progress was being made in both fields, Jerry Fodor and Zenon Pylyshyn—two function-first theorists—lodged a major complaint about connectionist architectures in 1988. Connectionism, they argued, failed to reproduce the fundamentally "systematic" nature of cognition (Symons & Calvo, 2014). To explain the term "systematic," they used a single sentence as an illustrative example: "Mary loves John." For a person to comprehend such a sentence, they argued, that person must also be capable of understanding the inverted form of the sentence: "John loves Mary" (Borensztajn et al., 2014).

To be useful, however, Fodor and Pylyshyn's quality of "systematicity" must be an explicitly defined category rather than just an example. Working with others, Dr. Gideon Borensztajn clarified Fodor and Pylyshyn's 1988 definition:

> A system is weakly systematic if it can generalize to novel sentences in which a word appears in a grammatical position it did not occupy in a training sentence but which it occupied in other training sentences. A system exhibits strong systematicity if, in addition, 'it can correctly process a variety of novel simple and novel embedded sentences containing previously learned words in positions where they *do not appear* in the training corpus . . .' (Borensztajn et al., 2014)

To summarize Borensztajn's definition, systematicity is the ability of a cognitive agent to abstract the semantic meaning of a word from its position in a specific grammatical context. Understanding the sentence "John loves Mary," for example, requires the ability to understand the sentence "Mary loves John." Understanding itself is dependent upon the meaning of the words "John," "Mary," and "love" in addition to their respective positions in the particular sentence.

The presence of systematicity within connectionist models has been a heavily debated topic since Fodor and Pylyshyn originally lodged their complaint. In their (2014) essay, John Symons and Paul Calvo reiterated this critique, arguing that connectionism "inevitably fails to provide a meaningful explanation of cognition insofar as it confuses the intrinsically systematic nature of thought with a system of associations." In other words, critics of the connectionist paradigm underscore the failure of association-based architectures, arguing that they cannot reproduce an important function of cognition: systematic thought. The next section—ranking thoughts—further details this argument.

The debate over systematicity clarifies the perspective of the function-first theorist. Rather than guess at neural architectures and observe the properties that emerge, function-first theorists want to begin by identifying all of the "functions" of the mind. The appeal of this approach has rallied scientists to expand and clarify the list of mental functions. As a result, systematicity is now little more than one function out of the multitude function-first theorists have identified.

Dealing with verbal descriptions of these functions can be quite cumbersome, however. To determine whether two experimental results evidence the same verbally-defined function, scientists would have to debate over the presence of the function in its totality. This obvious inefficiency has driven many function-first theorists to adopt a “symbolic” approach, a subset within the function-first paradigm that agrees upon one simple assumption for the sake of rapid progress. Symbolic theorists accept a correspondence between symbols—variables that can be manipulated in a mathematical scheme—and “cognitive representations.” The authors of *Categorizing Cognition* define cognitive representations as “internal states containing information that can be used by an animal, whether human or nonhuman, to interact with the environment in an adaptive manner” (2014, p. 27).

Later on, the same authors further explicate the correspondence between symbols and cognitive representations. For a particular example in their book, they define the symbols m_i and m_0 and relate them to one another in a cognitive process. Of these symbols they write, “The mental states m_i and m_0 may represent ideas (Locke, 1690/1924), things or objects (James, 1890), stimuli and responses (Hull, 1943), or events (Lieberman, 1993)” (2014, p. 39).

With citations that harken back to Locke’s work from the 1600s, it is clear the symbolic theorists do not demand precise formulation of *how* symbols represent cognitive representations. Rather, they argue, accepting this correspondence as fact allows for the efficient identification of mental functions.

A symbolic system, therefore, focuses on the functions of a desired cognitive architecture before determining the possible ways in which it could be implemented in the brain. However, this order of operations does not endow the first task with greater difficulty. If anything, scientists tackling the second face a greater challenge. The realistic neural implementation (biological or artificial) of symbolic architectures is not immediately clear. As a result, a multitude of potentially realistic implementations can result from a single symbolic architecture.

The authors of *Categorizing Cognition* synthesize previous such work done by symbolic theorists. To evaluate their success at doing so, the many functions explicated using symbols—including systematicity—must be replicated using a realistic neural architecture.

Ranking Thoughts

The backbone of *Categorizing Cognition* is the argument that many crucial mental functions become possible using working memory. This claim is based heavily upon Oberauer's "Design for a Working Memory," which outlines a possible implementation of cognition within working memory.

Another critical hypothesis outlined within *Categorizing Cognition* views symbolic process as more than a useful tool for analysis. Rather, the authors argue, the ability to manipulate symbols and grasp their correspondence with underlying cognitive representations is an essential feature of cognitive agents. Utilizing Oberauer's paper, they demonstrate a possible implementation of symbols within working memory.

In this regard, a symbol is "one building block of cognitive representation . . . symbols depend on a system of operations for recursively composing them into more complex symbols, which can be further composed, and so on" (2014, p. 29). Symbols, therefore, are the fundamental units of cognitive processing; their composition yields higher order cognitive representations. To make this argument precise, the authors of *Categorizing Cognition* formulate a ranking system into which various cognitive processes can be categorized.

The first rank—rank 0, or "nonstructured" processes—refers to simple relationships of varying strength between cognitive representations. As an example, the authors describe the relationship that can be formed between mental representations of rain and the wet ground:

> Perception of the external event, rain, produces a mental state ms_{rain}, representing rain, and perception of wet ground produces a mental state $ms_{wetground}$, representing wet ground. There is an associative link between ms_{rain} and $ms_{wetground}$, which is strengthened by contiguity of rain and wet ground and represents the statistical links in the environment, viz., that rain is likely to make the ground wet. (p. 38)

In symbolic language, the rank 0 relationship between thoughts of rain and the wet ground can be represented simply as $ms_{rain} \rightarrow ms_{wetground}$.

The limitations of rank 0 processes severely restrict the types of thought they can produce. Primarily, the hierarchical forms of thought required by systematicity, for example, cannot be accounted for using only rank 0 processes. The authors elaborate, "It is part of our formal definition of nonstructured processes that an associative link per se cannot be a component of another association: For example, we cannot have $(m_1 \rightarrow m_2) \rightarrow m_3$. By contrast, symbolically structured processes can be formed into recursive, hierarchical structures using higher-order relations" (p. 39).

The second rank of cognitive processing affords greater flexibility, but the limitations are still quite noticeable. Termed "functionally structured," rank 1 processes carry the ability to "reinterpret" cognitive representations. This ability is important, as representations do not always play the same role in different environments. The classic example of a problem unsolvable using rank 0 representations is the conditional discrimination problem, shown below. Within the problem, a cognitive agent is asked to learn the associations detailed in the "original task" box. Using a simple associative model, the solution would be impossible to obtain. Upon learning the first line, for example, black would become associated with the positive response, R+. Learning the second line would be impossible without somehow destroying the learning that occurred during the first association. By reinterpreting the inputs, however, rank 1 processes can compose more meaningful relationships between cognitive representations.

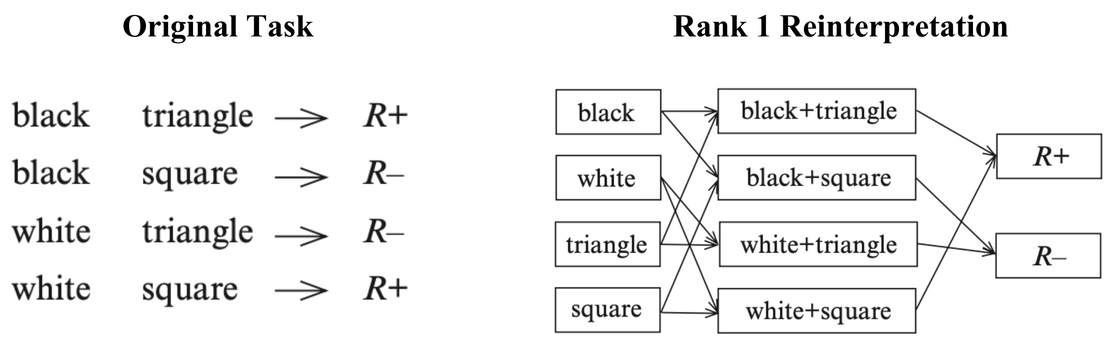


Figure 1: The conditional discrimination problem. The neural network on the right shows how representations can be "reinterpreted" to allow for more complex relationships. Adapted and reprinted from *Categorizing Cognition* (p. 37), by Halford et al. Copyright 2014 by The MIT Press.

Despite the enhanced capabilities of rank 1 processes, they are still limited. Higher cognition necessitates even higher ranks. *Categorizing Cognition* classifies all higher ranks into one category: symbolically structured processes, or ranks 2+. Unlike lower-rank processes, rank 2+ processes require "structural alignment." As defined by the authors, structural alignment entails "a binding of fillers to roles" (p. 46).

To illustrate this concept, the authors use the relation "largerThan" as an example. By using this relation, a cognitive agent can cognize a rather complex relationship between two other cognitive representations. For example, "largerThan(elephant, mouse)" is the symbolic representation of an agent's knowledge that an elephant is, in fact, larger than a mouse. The relation "largerThan" takes as its inputs two cognitive representations whose order matters. largerThan(elephant, mouse), they write, is very different than largerThan(mouse, elephant). In other words, symbolic relations specify “roles” which certain “filler” representations—the mouse and the elephant in this case—fill.

By understanding this simple example, it is possible to see how greater complexity is possible using rank 2+ processes. The relationships between cognitive representations no longer need to be purely association-based. They can be of higher ranks wherein more complex relationships are formed. The key to making this possible, the authors argue, is working memory. Working memory allows the brain to temporarily specify cognitive representations (fillers) to fill the roles required by a relation. The process of filling a role with a cognitive representation is termed “binding.” Working memory, they argue, allows for the transcendence of simple, associative processes by allowing for the implementation of roles, binding, and fillers.

Actual Entities

To apply Alfred North Whitehead’s process metaphysics to cognitive science, it is first necessary to introduce it. This section summarizes the fundamental tenets of Whitehead’s metaphysics, focusing primarily on their relevance to the study of cognition.

As it is a discipline with diverse approaches, metaphysics lacks an authoritative definition. In framing a metaphysical theory, it is the job of the philosopher to explicate his domain of study. Generally, metaphysical philosophy tends to study the nature of reality—the characteristics and structure of *everything*. It is therefore unclear what special place cognition holds in any metaphysical scheme. If a metaphysical theory is to be useful to the study of cognition, it should impart a better understanding of the characteristics of cognitive processes. Assuming cognition is no more than an example of a more general phenomenon within a metaphysical framework, that metaphysical framework does little to describe the particulars of cognition. Instead, it only describes the properties of the more general category to which cognition belongs. To the mathematician and philosopher Alfred North Whitehead, however, cognition lies near the apex of an ontological pyramid. Whitehead's analysis, therefore, is particularly valuable to cognitive science. This section intends to summarize this argument for the centrality of cognition, priming a discussion bridging his philosophy with cognitive science.

Alfred North Whitehead first developed his metaphysical framework, process metaphysics, in the 1920's. During this time period, Whitehead inked his ideas down in two volumes: *Science and the Modern World* and *Process and Reality*. Originally published in 1925 and 1929, respectively, the methods employed in each book reflect the relative stage of Whitehead's philosophical development. The philosopher Elizabeth Kraus eloquently contrasts the two books, writing, "[*Science and the Modern World*] represents Whitehead in his moment of romancing the metaphysical implications of his earlier epistemological theories, the full schematization [*Process and Reality*] having yet to be constructed" (1979, p. 11). *Process and Reality*, in other words, makes precise the metaphysical framework conceived of in *Science and the Modern World*.

In Whitehead's view, the domain of metaphysical study encompasses all *things* that are *real*. The goal of metaphysical study is to divide this "metaphysical domain" into a set of categories that are both mutually exclusive (disjoint) and collectively exhaustive (spanning). By accomplishing such a task, a metaphysician can create a minimal collection of terms by which any *thing* can be properly labelled and accounted for.

To understand Whitehead's categories (which he refers to as his "categoreal scheme"), it is necessary to first understand the domain that they partition. Whitehead circumscribes this domain using two terms: “actuality” and “experience.” Each of these terms refers to a property possessed by all the elements of the metaphysical domain.

Actuality, as defined by Whitehead, is the quality of concreteness. It is the property of real rather than theoretical existence. He first delineates this idea in *Science and the Modern World*, where he writes, “Nothing ever really recurs in exact detail. No two days are identical, no two winters. What has gone, has gone forever” (1925, p. 5). In this quotation, Whitehead equates actuality with specificity. Winter, to adopt his example, does not possess actuality. The winter of the year 1929, however, does. The former is a category, while the latter is an instance.

Whitehead castigates those who confuse mental objects of purely categorical existence with objects that possess actuality. Such transgressions commit the fallacy of misplaced concreteness, “the error of reifying what in fact is a high-order abstraction” (Kraus, 1979, p. 1). Actuality, therefore, is a requirement for some *thing* to qualify as a constituent of Whitehead’s metaphysical domain.

To better understand the characteristics of actuality, this section will briefly contrast Whitehead’s views on the metaphysical domain with those delineated by his predecessors. Although such views have mutated into a colorful variety, the paper will focus on the opposition of process metaphysics to any physicalist account of actuality. To properly describe physicalism, the philosopher Gregg Rosenberg provides a simple, yet quite lucid, definition: “Physicalism says that the fundamental physical facts are the only fundamental facts” (2004, p. 13). In other words, there are no incontrovertible *things* other than those describable by the laws of physics. On a purely intuitive level, the physical world seems to differ from that of the mental on a number of accounts. As the philosopher Galen Strawson describes, “what we ordinarily think of as the physical world [is] clouds, brains, chairs, mountains, and all the entities and qualities whose existence physics is right to recognize, quarks, say, or charge, or fields" (Strawson, 2019).

To Whitehead, physicalists regularly fall victim to the fallacy of misplaced concreteness. By restricting the metaphysical domain to a world composed of well-known objects, physicalists discuss abstractions as if they possess intrinsic actuality. In contrast, Whitehead outlines a drastically different metaphysical domain when he explicates the purpose of his metaphysics, which he refers to as "Speculative Philosophy":

> Speculative Philosophy is the endeavour to frame a coherent, logical, necessary system of general ideas in terms of which every element of our experience can be interpreted . . . everything of which we are conscious, as enjoyed, perceived, willed, or thought, shall have the character of a particular instance of the general scheme. (1929/1978, p. 3)

By making such a claim, Whitehead became an early adopter of "panexperientialism," the view that only those *things* possessing experience can be labeled as actual. Later in *Process and Reality*, he argues, "The elucidation of immediate experience is the sole justification for any thought" (1929/1978, p. 4). Without the presence of experience, he argues, there is no *thing* to which you can refer.

In summary, Whitehead views the metaphysical domain as a set of *things*, all of which possess both actuality and experience. Having circumscribed his beliefs on the metaphysical domain, Whitehead focuses the remainder of his philosophical efforts towards the characterization of each *thing* within the metaphysical domain.

It seems self-evident, Whitehead argues, that the division of an experience yields something incomplete. Rather, experiences are composite units—"drops of experience," as Whitehead refers to them. By removing part of a drop of experience, its composite unity must be destroyed in the process. This is not to say that an experience is indivisible—rather, it is to say that the division of a drop of experience cannot possibly yield another drop of experience.

Whitehead, equating experience with actuality, terms these composite instances of experience “actual entities.” Of these actual entities he writes:

> [Actual entities] are the final real things of which the world is made up . . . God is an actual entity, and so is the most trivial puff of existence in far-off empty space. But, though there are gradations of importance, and diversities of function, yet in the principles which actuality exemplifies all are on the same level. The final facts are, all alike, actual entities; and these actual entities are drops of experience, complex and interdependent. (1929/1978, p. 18)

Actual entities are therefore the *things* that collectively make up the entire metaphysical domain. All of them share two fundamental characteristics: actuality and composite unity of experience.

Whitehead contrasts the words “experience” and “mentality,” although their difference is initially unclear. Inevitably, this lack of clarity is a result of humanity’s inability to divorce the two—a human’s subjective experience is inseparably defined by both (Kraus, 1979). To elucidate this distinction, the paper will refer to the work of Gregg Rosenberg, a philosopher who has clarified many of Whitehead’s theories on panexperientialism. Rosenberg utilizes the term “field of experience” to mean a grouping of subjective qualities within an experience. He writes the following of these fields:

> The best term for the alien character of these fields is *protoconscious*, a term meant to suggest that they contain experienced qualitative objects that are not, strictly speaking, being experienced by a mind (because there is no associated cognition). These

protoconscious states are states of pure experience. They need not have semantic content, and certainly no cognition will occur within the manifold of experience. (2004, p. 94)

From this description, it seems that Rosenberg is attempting to distance the concept of experience from those of cognition, consciousness, and mentality. As discussed in the previous section, these three concepts are not necessarily synonymous. Each does, however, require experience to exist.

This is the fundamental relationship between cognition and experience, as defined by Alfred North Whitehead. Cognition, as a term, defines a more stringent set of requirements than does its parent category (1929/1978). In this sense, the results of cognitive processing—cognitive representations—collectively constitute a subset of the universal to which they belong: that of the actual entity.

Throughout *Science and the Modern World* and *Process and Reality*, Whitehead endows actual entities with a number of properties. As an example of an actual entity, cognitive representations should exhibit all of these properties and more.

Genetic Analysis

The popularity of the physicalist paradigm begs the question: why is the fallacy of misplaced concreteness—the mistake of confusing physical objects with actual entities—so alluring? Clearly there must be some quality about the human experience that suggests the actuality of chairs and mountains and all the other physical objects listed by Strawson. Throughout *Science and the Modern World*, Whitehead (1925) repeatedly labels this paradigm the "new mentality," a fallacious mode of thought that mistakes mental representations of objects for some cohesive physical constitution.

It is obvious that the concept of a chair is real insofar as the concept imparts a real difference on a particular experience. Within the confines of that experience the chair exists as an apparently cohesive unit. However, the mental representation of the chair is a different entity from the multiplicity of data that give rise to it. These data might include, for example, the visual images of the legs and seat, the tactile feeling of the wood, and memories of other chairs encountered in the past.

Whitehead (1929/1978) suggests that there exists a "vector character" of these mental representations—they point towards something else. However, unlike his predecessors, many of whom believed in the fundamentally different constitutions of mental representations and referent data (substance dualism, for example), Whitehead underscored their common actuality. For an object or data point to exist—that is, for it to be an element within the metaphysical domain—it must be an actual entity.

The vector character of actual entities towards other actual entities begs the natural question: "How can concrete fact exhibit entities abstract from itself and yet participated in by its own nature?" (Whitehead, 1929/1978, p. 20). In other words, how do actual entities possess vector character? Applied to the field of cognitive science, this question can be reformulated: how can symbols represent cognitive representations during a symbolic process? Whitehead answers his own question with a detailed scheme elucidating the ways in which actual entities ingress other actual entities into their constitution. He names this scheme the "genetic analysis" of an actual entity, the study of an actual entity's creation—its genesis (1929/1978).

Process

One of the most fundamental qualities of an actual entity is its composite unity. To summarize arguments from previous sections, an actual entity has a composite unity of feeling because it is a “drop of experience” whose subunits lack full constitution. However actual entities also possess vector character—they refer to other actual entities. The process by which an actual entity is created, therefore, begins with data elsewhere in the world (the data to which the vector in “vector character” points) and ends in the composite unity—the fully unified endpoint that possesses a singular feeling. Whitehead (1929/1978) refers to this process simply as “process,” the root term in the phrase “process metaphysics.” Properly surveying the intricacies of process is the task of books, not undergraduate theses. For this reason, this section will summarize the relevant details of process and leave the reader to explore the subject further if he or she is so inclined.

Process is most easily understood as a sequence of events, each of which yields a product that more closely resembles the composite unity. However, the interpretation of this “sequence” as temporal misrepresents the nature of actual entities. For an actual entity to possess a composite unity, it must not be composed additively of diverse elements and feelings. Rather, each component of an actual entity must play a specific role in the composite unity.

Dr. Kraus (1979) summarizes this property, writing that elements within an actual entity can be separated by examining the ways in which they contribute to the composite unity. Remember that subsets of an actual entity cannot be “actual” in isolation. Similarly, elements within the “sequence” of a process do not produce anything actual on their own. Composite unity is only possible when *all* of the elements within a process participate. The elements of this sequence, therefore, occur simultaneously. Despite this fact, however, it is possible to analyze an element of this sequence by examining the ways in which it influences the composite unity.

This section will focus on one particular event within process: the transition from the initial data to the objective datum. The initial data are (collectively) all of the data that an actual entity *can* access. They are the actual entities to which the vector in "vector character" points. The objective datum constitutes the subset of those data which the process incorporates into its satisfaction (Whitehead, 1929/1978; Kraus, 1979).

Below is a useful diagram that illustrates this transition in simplistic terms. In addition to the formation of the objective datum, a multitude of additional steps carry the actual entity to satisfaction.

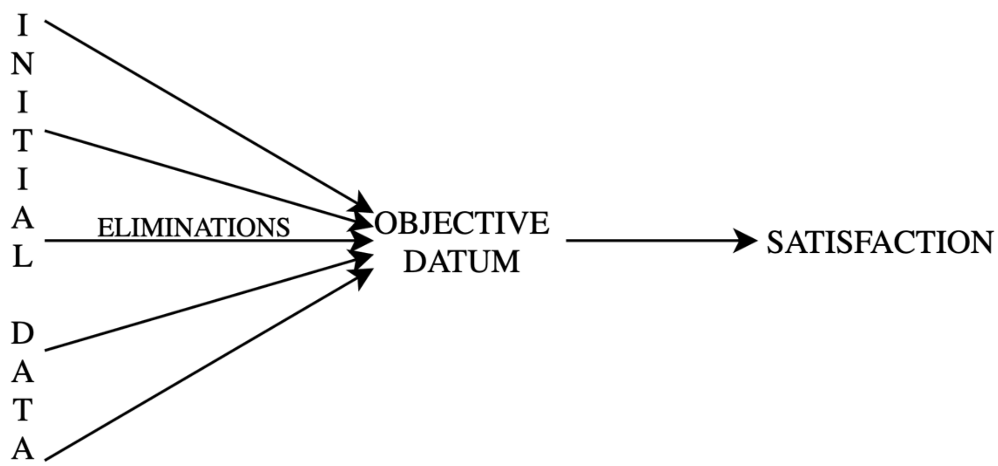


Figure 2: The transition from initial data to objective datum requires eliminations. Subsequent steps carry the objective datum to satisfaction. Adapted from *The Metaphysics of Experience: A Companion to Whitehead's Process and Reality* (p. 118), by E. Kraus, 1979. Copyright 1979 by the Fordham University Press.

To fully summarize this fraction of process, it is necessary to detail each subsidiary component. As explained by Whitehead, the initial data are a "multiplicity," while the objective datum is a "nexus" (1929/1978).

To define these terms, Whitehead tells his readers simply that "a multiplicity consists of many entities" (1929/1978, p. 24). A multiplicity, in other words, is no more than a collection of actual entities. Recall that actuality, a property of all *things* within the metaphysical domain, requires a composite unity of experience. The collection of entities itself possesses no actuality because it has no composite unity. Rather, a multiplicity only becomes actual insofar as it participates in a particular process.

To understand why the initial data are a "multiplicity" of data rather than a single datum or no data at all, it is only necessary to refer back to Whitehead's earlier axioms. To begin, Whitehead states that actual entities possess vector character. The initial data, therefore, cannot be an empty set. Furthermore, the composite unity of an actual entity requires that the initial data be composed of more than one entity. As Dr. Kraus explains, "The total satisfaction of one entity cannot enter into the composition of another entity in its full complex unity without that entity's quite literally *becoming* the entity absorbed" (Kraus, 1979, p. 103). In other words, because no part of an actual entity can be eliminated without having that entity lose its composite character, a process with a singular actual entity as its initial data would simply result in that actual entity. The term "multiplicity" therefore makes sense, as process requires multiple entities for a meaningful starting point.

After being provided with the proper set of initial data, a process must begin its transition towards composite unity. To do this, the actual entity eliminates elements within each initial datum disagreeable to the initial data's constitution as a whole. Such eliminations are most easily understood in the language of prehensions. As Whitehead explains, "the negative prehensions which consist of exclusions from contribution to the concrescence can be treated in their subordination to the positive prehensions. These positive prehensions are termed 'feelings'" (1929/1978). As Kraus explained above, elements are not incorporated into the constitution of an actual entity additively. Rather, they all play a role in the composite unity of the actual entity. The perspective of an actual entity on each of these elements is termed "prehension." In other words, the way that an actual entity incorporates another actual entity into its composite unity is how it "prehends" that actual entity.

Prehensions can be divided into two types: positive and negative. A negative prehension occurs when an actual entity simply excludes part of another actual entity from any participation in its satisfaction. A positive prehension is the opposite: an actual entity somehow incorporates data from another actual entity into its constitution. In this sense, the positively prehended actual entity transfers part of its composite unity feeling to the actual entity undergoing process. This is why a positive prehension is referred to as simply "feeling."

As detailed above, entire actual entities cannot be positively prehended, as prehensions are supposed to constitute only partial elements in the composite unity. Instead, prehensions function by eliminating elements of an initial datum which will not play a role in the composite unity. Whitehead proceeds to explain how prehension yields entities lacking composite character:

> A prehension reproduces in itself the general characteristics of an actual entity . . . In fact, any characteristic of an actual entity is reproduced in a prehension. It might have been a complete actuality; but, by reason of a certain incomplete partiality, a prehension is only a subordinate element in an actual entity. (1929/1978, p. 19)

This partiality is an obvious consequence of the elimination involved in a prehension. Because elements of the object of prehension are eliminated, the product of prehension will no longer possess composite unity.

This further demonstrates the necessity of a multiplicity of initial data. The eliminated portions of an entity within the initial data are supplanted by portions from other entities within the initial data. It is through this process that fractional portions of each initial datum are synthesized to form the objective datum. As a result, the nexus constituting the objective datum is an “operation of mutually adjusted abstraction” (Whitehead, 1929/1978, p. 210). Eliminations occur so that the elements of the objective datum are in agreement with one another. No two feelings can contradict each other within the objective datum.

Although Whitehead expands upon these ideas to detail other intricacies involved in the creation of an actual entity, the transition of the initial data to the objective datum is the information most relevant to the argument in the paper. For this reason, no more detail regarding Whitehead’s genetic analysis is necessary.

Symbolic Accounts of Subject

As argued previously, cognitive representations of any rank exhibit many of the same properties as actual entities. They should, therefore, be analyzable in the same way as Whitehead’s genetic analysis. Much like actual entities, they should begin with a multiplicity of initial data, synthesize this data into a singular objective datum, and conclude with a composite unity of feeling. This section begins by identifying the many ways in which cognitive representations properly display the features of Whitehead’s genetic analysis. In contrast, this section also identifies the ways in which symbolic process deviates from Whitehead’s genetic theory. Finally, these deviations serve as red flags for the easy identification of defects within *Categorizing Cognition* and “Design for a Working Memory.”

The authors of *Categorizing Cognition* lean on a dual-system theory of the brain. The first system contains modules that are highly specialized for given tasks. A module is, in their own words, a “fast, domain-specific set of processes that are informationally encapsulated, impose only very low processing loads, are not strategically modifiable, and show little or no relationship to individual differences in intelligence” (2014, p. 31). One prime example of a modular system is the visual system, which is rapid yet inflexible. They proceed to contend that "many modular processes are not symbolic and do not entail explicit, accessible representations of relations" (2014, p. 32). Clearly, modules within the brain cannot account for all mental phenomena. Particularly, the rapidity of calculations within a module of the brain does not seem to explain the contents of our experience. Cognition feels like a slow, thought through process with linear flow rather than hyper-parallelism.

To explain this dissonance, *Categorizing Cognition* refers to the model described in "Design for a Working Memory." To Oberauer, working memory is the key to cognition. It provides a neural substrate within which cognitive representations can be "bound" to their roles in a symbolic process. He writes, "These bindings must be dynamic, which means that they can be set up quickly and dissolved quickly when the structure is updated or discarded" (2009, p. 47).

Unlike neural modules, working memory is tied closely with attention. One popular view contends that working memory is the region of the brain giving rise to our notion of attention (Oberauer, 2019). Regardless, the contents of our experience—the things which we are "attending" to—seem to be the contents of working memory. This identification is one way in which the working memory theory fulfills the obligations of proper genetic analysis. If symbolic processes individually dominate working memory, only to be dissolved and replaced by subsequent symbolic processes, their composite unity seems to be a natural byproduct.

Symbolic processes also begin with a multiplicity of initial data. Take, for example, *Categorizing Cognition*'s example of feeding animals. The idea "Jane feeds cat" can be represented using the symbolic composition "feeds(Jane, cat)." This obviously conveys a very different meaning from the sentence "Cat feeds Jane," which would result in the symbolic composition "feeds(cat, Jane)." Working memory discerns between the two using dynamic bindings, linking cognitive representations to roles in the symbolic process. In this instance, the roles take the form of verb(subject, object). Each synthesis, therefore, seems to begin with multiple cognitive representations (the multiplicity) and end in a composite unity (the symbolic representation).

Symbolic models of cognition diverge from Whiteheadian process over the topic of feelings. For the purpose of the paper, Whitehead's term "feeling" will be used interchangeably with the cognitive science term "meaning." Recall that *Categorizing Cognition* defines cognitive representations using footnotes to Locke. Within *Process and Reality*, Whitehead describes process metaphysics as a further refinement of Locke's original philosophy. *Process and Reality* and *Categorizing Cognition* therefore seem to be discussing the same concept—the idea of a concept itself. The words "meaning" and "feeling" seem to refer to the same thing: the subjective quality of the cognitive representation.

Problems arise over "feelings" because it is not entirely clear how representations of high rank incorporate feelings from perceptual inputs. In previous criticisms of symbolic theories of mind, this issue has been termed the "symbol grounding problem." How do symbols adequately convey the meaning of their referent cognitive representations? As Stevan Harnad notes in his (1990) essay concerning the symbol grounding problem, "The standard reply of the symbolist . . . is that the meaning of the symbols comes from connecting the symbol system to the world 'in the right way'. But it seems apparent that the problem of connecting up with the world in the right way is virtually coextensive with the problem of cognition itself" (p. 340). Harnad's question closely parallels Whitehead's original question from 1929: how do actual entities possess vector character? To phrase the question in multidisciplinary language, how does a symbolic composition utilize the feelings or meanings of its initial data?

The authors of *Categorizing Cognition* attempt to dilute the worries of symbolic criticisms by writing:

> Structural correspondence by itself would not be enough because of the symbol grounding problem, and the links between mental states and external entities . . . are necessary to keep representations anchored in reality. Thus, the symbolically structured level complements, rather than displaces, the functionally structured level of processing. (2014, p. 50)

This quotation contains the bulk of the symbolic theorists' response to Harnad's earlier criticism. The meaning or feeling of a high-rank thought, they argue, cannot be expressed by the symbolic expression (feeds(cat,Jane), for example) alone. Rather, meaning is also sourced from the functionally structured level (the representations referred to by the symbols *cat* and *Jane*—the initial data). By maintaining the neural activation of these initial data, the symbolic synthesis somehow transforms their feelings into a composite unity of feeling.

Although Oberauer's model preceded *Categorizing Cognition* by five years, his explanation was largely the same. Within his (2009) paper, Oberauer explicates the ways in which representations are saved to long term memory for easy retrieval. He coins the term "chunking" to refer to the encapsulation of a relational representation:

> A chunk is a representational unit in which other units and their relations are packed so that they are not individually accessible—unless the chunk is unpacked again … Upon encountering for the first time an episode in which the pastor calms the businessman, the system must create a new chunk representing that proposition. The proposition chunk

> would be associated with the representations of the three concepts involved (i.e., pastor, calming, and businessman). In addition, it needs to be associated with representations that code the conjunction of each concept with its role in the proposition (i.e., the conjunction of “pastor” with agent, of “businessman” with patient, and of “calming” with action). (2009, p. 78)

In summary, the results of symbolic processes are saved in a way that allows them to be reloaded into working memory. During this reloading process, the symbolic representation is reconstructed and the initial data (the representations referred to by *cat* and *Jane*, for example) are reactivated. Only by doing this can the meaning of the memorized representation be accessed again.

Naturally, the model twice outlined by *Categorizing Cognition* and “Design for a Working Memory” seems to imply that there are two sorts of representations. The first type is sourced from the modular inputs discussed earlier. Representations of this type arise through perception. To qualify as a perceptual representation, the mental object must simply possess “structural correspondence between the representation and the represented structure” (Halford et al., 2014, p. 29). Here, the represented structure can be as simple as sensory data. In this way, perceptual cognitive representations refer to the world external to the cognitive agent. Such representations include, for example, the visual image of a chair, the sound of a nearby tree falling, or the ideas that the symbols “Jane” and “cat” refer to. These data form a baseline of worldly objects about which the cognitive agent can learn.

In contrast with perceptual representations, representations of the second type are the results of processing that occurs within working memory. Effectually, they are relational representations because they relate perceptual representations to one another. These relational representations—in the same way as perceptual representations—can participate in symbolic processes as initial data by binding to roles. The authors of *Categorizing Cognition* detail, "Symbolization enables relational representations to be arguments to other representations, yielding higher-order relations" (2014, p. 66). The fundamental hypothesis outlined in *Categorizing Cognition* asserts the necessity of such hierarchical encapsulation.

The real division between perceptual and relational representations, however, is the presence of intrinsic meaning. The symbols "elephant" and "mouse" in the relational representation "largerThan(elephant, mouse)" refer to *meaningful* perceptual representations. The authors of *Categorizing Cognition* make the implicit assumption that certain representations possess this sort of intrinsic meaning, making the symbol grounding problem a nonissue. In contrast, the relational representation—symbolically denoted "largerThan(elephant, mouse)" in this case—is dependent upon the participating perceptual representations for meaning. Relational representations therefore possess extrinsic meaning.

Although the transition from perceptual representation to relational representation appears to mirror the process metaphysical transition from the initial data to the objective datum, symbolic frameworks violate the key principles Whitehead lays out in his genetic analysis. Namely, symbolic processes of this sort conjure no more than a multiplicity, leaving the result devoid of composite unity. For a relational representation between three symbols to possess meaning, for example, the three referent perceptual representations must be kept active. The feelings included in these perceptual representations are wholly included in the feelings of the derivate relational representation. The additive nature of such a transference of feeling renders relational representations stuck in the phase of the multiplicity.

Within Whiteheadian process, positive prehensions constitute an incorporation of certain feelings within one actual entity into the constitution of a subsequent actual entity. Whitehead asserts that, to create some sort of composite unity, the data composing the prehended actual entity must be only partially retained. Additionally, these feelings must be understood in accordance with the composite unity. For a symbolic process to result in a composite unity of feeling, it too must perform eliminations on its initial data. Only by doing so can symbolic processes produce meaningful ideas rather than memorized relationships.

In summary, the genetic analysis of symbolic processing has resulted in the paper's first concrete suggestion for future study. To create a composite unity of feeling, each symbol used to create a relational representation must somehow propagate only a subset of its referent perceptual representation's feelings. Additionally, the symbolic process must be able to coordinate these subsets into a cohesive whole. By explaining the methods by which these phenomena are accomplished, symbolic theorists may be able to create a viable account of cognition.

Coordinate Analysis

Within *Process and Reality*, Whitehead defines two ways in which an actual entity can be considered. The first, as discussed within the last section, is genetic analysis—the study of the method by which an actual entity comes into existence. Genetic analysis constitutes a division of the actual entity into elements that each explain one aspect of process. The second mode of analysis, in contrast, examines an actual entity once it has fully achieved its composite unity. This type of analysis, which Whitehead terms "coordinate analysis," is little more than a different perspective on the same phenomena. Both halves of *Process and Reality* study the interplay between actual entities, but each half examines a portion of process in greater detail.

To parallel the paper's section on genetic analysis, this section will begin with Whitehead's delineation of coordinate analysis and proceed to apply it to "Design for a Working Memory" and *Categorizing Cognition*.

The Extensive Continuum

Whitehead's coordinate analysis (and his metaphysics as a whole) refers heavily to the idea of the "extensive continuum," a concept analogous to Einsteinian spacetime. In putting forth his theory of the extensive continuum, Whitehead argues that the most descriptive topology of reality looks very different from frameworks elucidated in physics. Rather, it is composed of regions with the potential to form an actual entity. Once an actual entity achieves a composite unity within a given region, it "atomizes" that region. To elaborate, Kraus writes:

> [The] spatial region actualized in a [process] … is an indivisible unit of real space, interlocked with all other regions because of its actualization of the general schematic relations of the extensive continuum it atomizes … [In] any given spatial extensity, not all possible perspectives must be realized. (1979, p. 127)

In other words, an extensive continuum contains all of the possible forms that a particular actual entity could have assumed. This scope of possibility is a continuum (a spatial region), allowing for infinitesimal differences in quality. When an actual entity achieves satisfaction, its determinate existence makes only one location within the continuum actual. The rest of that particular continuum is merely what could have been.

The addition of coordinate analysis produces insight obscured during genetic analysis alone. Namely, the ways in which an actual entity intervenes "into processes transcendent to itself" becomes clear (Kraus, 1979, p. 105). An actual entity, behaving in this way, acts as "superject." Whitehead contrasts this term with an actual entity acting as "subject," the behavior of an actual entity as it is created. An actual entity acts as subject during its own process, whereas an actual entity acts as superject when it is an initial datum for a future process.

It is easy to oversimplify the process by which an actual entity acts as superject. Due to the composite unity of actual entities, it is easy to view the initial data as objects whose entire constitutions are incorporated into the actual entity acting as subject. This description, however, ignores the fact that eliminations must occur for a novel actual entity to form. To determine the components of an initial datum to eliminate, process involves a genetic analysis of that initial datum. In this regard, genetic analysis is not simply a philosophical tool—it is a phenomenon that actually occurs in the world (Kraus, 1979).

It is by this thought process that Whitehead draws his fundamental conclusion about the existence of an actual entity: it exists as nothing more than the process by which it came into existence. The individual process creating an actual entity *is* that actual entity. It must follow that, when an actual entity acts as superject, it is the components of its process that intervene in a future actual entity. In this regard, the actual entity acting as subject and the actual entity acting as superject are identifiable.

Furthermore, since an extensive continuum contains the potential forms of an entity prior to actualization, it is a collection of the possible processes which could occur. These extensive continua are "interlocking," according to Kraus. Using a reversion of the ideas outlined during the genetic analysis, it is possible to see how this is true. Genetic analysis yields a given perspective on actual entities. This perspective shows a set of initial data given to a process to be reduced to a composite unity of feeling. This "process" gradually eliminates data and reinterprets feelings so as to endow the actual entity with composite unity. Coordinate analysis, conversely, yields the opposite perspective. Rather than view the process as it is created, it examines the extensive continuum that the process atomizes, complete with the atomized locus and the rest of the continuum existing only as what could have been.

By the same logic used in the genetic analysis, it is possible to see how these extensive continua interlock with one another. This interlocking character is best understood in the mathematical language of set theory. Each actual entity is a set of elements that collectively define its process of creation: the initial data used, the eliminations that occur, and other elements not detailed in the paper (Kraus, 1979). The interlocking relationship between two actual entities is not physically realized in any way. Rather, it is a product of the joint analysis of two actual entities. Two actual entities could be labelled disjoint with respect to one another, for example, if they shared no elements of process whatsoever. If they shared an initial datum, however, they would not be entirely disjoint.

Symbolic Accounts of Superject

Whitehead's coordinate analysis is no more than a different perspective on the same material studied by genetic analysis. This new perspective, however, primes a number of novel intuitions about the nature of cognition.

To reiterate one of Whitehead's key findings in coordinate analysis, an actual entity is no more than its process of creation. In this regard, the actual entity acting as superject—that is, the actual entity's "intervention into processes transcendent to itself" (Kraus, 1979, p. 105)—is identifiable with that actual entity acting as subject—the actual entity as it is being created. This subject-superject complex fully describes all the meaningful details about an actual entity.

The genetic analysis section of the paper argued that relational representations contain a multiplicity of feelings rather than a composite unity. Although the issues presented by such a system are not immediately apparent, when discussing the relationship between a relational representation and its constituent perceptual representations, it makes the encapsulation of symbolic representations problematic.

As knowledge grows hierarchically, relational representations of greater complexity form. However, as the symbolic theorists themselves argue, the feeling of a relational representation depends upon all the feelings of its constituent perceptual representations (Harnad et al., 2014). What happens, then, when a relational representation acts as superject in a more complex relational representation? To endow this more complex representation with meaning or feeling, the symbolic architecture would need to maintain the activation of *all* the subsidiary perceptual representations (Halford et al., 2014). For example, a relational representation composed from three other relational representations would need to derive its meaning additively from the meanings endowed to each subsidiary relational representation. Assuming each subsidiary relational representation derives its meaning from three perceptual representations, the relational representation being composed would derive its meaning from *nine* perceptual representations.

As the hierarchy of knowledge grows, the cognitive agent inevitably encounters what this paper will dub the "infinite queuing problem," wherein an immense number of perceptual representations must remain active for the resultant relational representation to possess feeling.

Obviously, the human brain does not encounter such an infinite queuing problem. Rather, highly complex relational representations seem to express their own composite unities of feeling about the same worldly data upon which perceptual representations are based. The authors of

Categorizing Cognition corroborate this fact by attempting to tie the meaning of a relational representation to the world. However, the assumption that relational representations simply collect ever-larger groups of perceptual representations is likely a false one.

Whitehead's solution to this problem takes two homologous forms. His first suggestion, as outlined within his genetic analysis, requires the elimination of data within each constituent perceptual representation so that the relational representation has composite unity (1929/1978). In this regard, the relational representation no longer expresses a sum of the feelings of its subsidiary perceptual representations. Rather, it expresses its own composite unity of feeling. Building off this claim, with the knowledge that an actual entity acting as superject has the same constituency as the same entity acting as subject, Whitehead's idea of extensive continua can be applied to the field of cognitive science.

For relational representations to behave like actual entities, they must "interlock" in the same way as extensive continua. In the context of symbolic systems, this implies that a relational representation can efficiently queue the set of feelings that compose its composite unity. *Categorizing Cognition* tries to achieve this goal by maintaining the activation of perceptual representations when a relational representation is present in working memory. As discussed in the previous section, however, these feelings remain in the state of the multiplicity. It is not clear how symbolic processes coordinate pertinent feelings within subsidiary representations in accordance with its composite unity. In other words, symbolic processes only possess the capability to memorize relationships between predetermined perceptual representations. They fail to describe how the feelings of the perceptual representations play a useful role in the relational representation.

Such a structure has made robotic memory architectures the forte of symbolic theorists. The most obvious example is that dually outlined by *Categorizing Cognition* and Dr. Oberauer. Chunks, Oberauer elaborates, are added to an ever-growing web of relational representations. Relational representations are connected to one another and to perceptual representations in long term memory using simplistic associations (Halford et al., 2014; Oberauer, 2009). In this manner, the symbolic agent can queue up "relevant" relational representations, where relevance is determined by associations with the current contents of working memory. This model, however, seems quite mechanistic. It reduces the human to a memorization machine rather than a cognitive agent capable of analyzing *why* the learned relationships between perceptual representations exist.

Older symbolic theories encountered similar difficulties. Gary Marcus, an artificial intelligence researcher and cognitive scientist, outlined a symbolic theory of cognition in 2001. He later revised a key principle of this theory, namely that "the mind has a neurally realized way of representing arbitrary trees, such as the syntactic trees commonly found in linguistics" (2014). Essentially, he had theorized that humans learn vast graphs (or trees) of relationships between words, such as the sort used for diagramming sentences. He revised this claim, instead asserting, "people don't behave as if they really can represent full trees (Marcus 2013). We humans have trouble remembering sentences verbatim (Qarvella 1971; Lombardi and Potter 1992); we have enormous difficulty in properly parsing center-embedded embedded sentences (*the cat the rat the mouse chased bit died*)" (2014).

Perhaps the symbolic theorists' vision of human memory architectures is wrong. The application of Whiteheadian philosophy insists upon the interlocking character of extensive continua. What would this look like in the context of perceptual and relational representations?

According to the genetic analysis of symbolic representations detailed earlier, it is clear that relational representations must eliminate and combine feelings from disparate perceptual representations. This novel composite unity of feelings can be described genetically by examining the feelings eliminated from each initial datum. Additionally, it can be described coordinately. The relational representation obviously shares a subset of the feelings of each of its perceptual representations. The relational representation also adds onto this subset, incorporating feelings from other perceptual representations.

By this logic, both relational and perceptual representations are describable by referencing the feelings they possess. Since, according to the authors of *Categorizing Cognition*, intrinsic meaning is endowed to a particular neural substrate—not working memory—the feelings in this substrate must interlock in the same manner as Whitehead's extensive continua. For example, two representations (of either type) could be easily related to one another simply by comparing their activation patterns within the neural substrate. Such a comparison would account for the complexities of the representations rather than their memorized symbolic constructors.

A framework based on Whitehead's extensive continuum would blur the distinction between relational and perceptual representations. Both types would exist as vehicles for labelling specific composite unities of feeling within the same neural substrate. Recall that eliminations must occur for an initial datum to participate in the composite unity of an actual entity (Whitehead, 1929/1978). For a perceptual representation to act as an initial datum, therefore, it cannot possibly be an indivisible, primordial entity like symbolic theorists implicitly assume. Rather, it too must label a composite (yet divisible) unity of feeling. Even though the perceptual representation may not be symbolically structured like a relational representation, it too must be learned.

This line of reasoning concludes with the second key suggestion for cognitive scientists. Rather than save a relational representation by explicitly storing the symbolic structure that creates it (Oberauer's theory of chunking), long term memory should be able to queue its composite unity of feeling. Effectively, such an architecture would interlock both relational and perceptual representations within the same neural substrate, allowing for a more natural, less mechanistic model of memory and mind.

Conclusion

Analysis of symbolic processes through a Whiteheadian lens yields intriguing results. Not only is it clear that symbolic theorists have yet to find an adequate neural implementation of their theories, but several directions for future study are easily identifiable.

Primarily, Whitehead offers a unique perspective on the symbolic grounding problem, highlighting the ways in which symbolic processes fail to represent the complexities involved in a cognitive experience. The symbolic paradigm's current treatment of perceptual representations, as practically indivisible entities with intrinsic meaning, yields a mechanistic cognitive architecture. Instead, the paper contends that perceptual and relational representations alike are only meaningful insofar as they label and queue a certain set of feelings within the same neural substrate.

Building on this conclusion, the paper suggests two concrete solutions. First, cognitive scientists should ensure that the relational representations they discuss are endowed with composite unity of feeling, and their models should involve symbolic processes that eliminate portions of each initial datum. Second, cognitive scientists should question the storage of relational representations in an area entirely separate from perceptual representations. By accounting for the interlocking character of relational representations within the same neural substrate as perceptual representations, symbolic theorists might be able to create a long term memory architecture with the properties of Whitehead's extensive continuum.

This is not a tacit rejection of symbolic theories of mind. The ranking system, for example, carries great value as it provides a strong rejection of association-based architectures. It is clear, however, that symbolic processes do not capture the entire truth of cognition. By incorporating ideas from Whitehead's process metaphysics, symbolic theorists may be able to narrow the gap between theory and reality.

References

Borensztajn, G., Zuidema, W., & Bechtel, W. (2014). Systematicity and the need for encapsulated representations. In Symons, J. and Calvo, P. (Ed.), *The architecture of cognition*. Cambridge, MA: The MIT Press. Retrieved from Amazon.com.

Gayon, J. (2016). From Mendel to epigenetics: History of genetics. *Comptes rendus biologies*, 339(7-8), 225-230. doi:10.1016/j.crvi.2016.05.009

Griffiths, T. L. et al. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Science*, 14(8), 357-64. doi:10.1016/j.tics.2010.05.004

Halford, G. S. et al. (2014). *Categorizing cognition: Toward conceptual coherence in the foundations of psychology*. Cambridge, MA: MIT University Press.

Hardcastle, V. (1996). *How to build a theory in cognitive science.*

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3), 335-346. doi:10.1016/0167-2789(90)90087-6

Kraus, E. M. (1979). *The metaphysics of experience: A companion to Whitehead's process and reality*. New York, NY: Fordham University Press.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Marcus, G. (2014). PDP and symbol manipulation: What's been learned since 1986? In Symons, J. and Calvo, P. (Ed.), *The architecture of cognition*. Cambridge, MA: The MIT Press. Retrieved from Amazon.com.

Miller, G. (2003). The cognitive revolution: a historical perspective. *TRENDS in Cognitive Sciences*, 7(3), 141-144. doi:10.1016/S1364-6613(03)00029-9

Murphy, G. L. (2019). On Fodor's first law of the nonexistence of cognitive science. *Cognitive Science*, 43: e12735. doi:10.1111/cogs.12735

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation: Vol 51* (pp. 45-100). Elsevier Academic Press. doi:10.1016/S0079-7421(09)51002-X

Oberauer, K. and Hein, L. (2012). Attention to information in working memory. *Current directions in psychological science*, 21(3), 164-169. doi: 10.1177/0963721412444727

Strawson, G. (2019) What does physicalist mean? A prolegomenon to physicalist panpsychism. In W. Seager (Ed.), *The Routledge handbook of panpsychism*. Oxford, UK: Taylor & Francis.

Symons, J. & Calvo, P. (2014). Systematicity: an overview. In Symons, J. and Calvo, P. (Ed.), *The architecture of cognition*. Cambridge, MA: The MIT Press. Retrieved from Amazon.com.

Rosenberg, G. (2004). *A place for consciousness: Probing the deep structure of the natural world.* Oxford, UK: Oxford University Press.

Whitehead, A. N. (1978) *Process and reality*. New York, NY: The Free Press. (Original work published in 1929).

Whitehead, A. N. (1925). *Science and the modern world*. New York, NY: The Macmillan Company.

Author Biography

Jake Sansom studied computational engineering at the University of Texas. He was originally born in Austin, but he grew up in Birmingham, Alabama. In addition to taking classes, Jake has spent much of his time working as a research assistant for the Willerson Center for Cardiovascular Modeling and Simulation, where he has developed computational models and open-source software for the analysis of experimental data. He has also enjoyed his time volunteering at the Rosedale School and the Arc, both of which provide support to people with cognitive disabilities. After graduation, Jake will be moving to San Diego to begin a systems engineering career as part of Northrop Grumman's rotational program. If the allure of research proves strong enough, however, he may redirect his plans towards the pursuit of a doctorate degree.