Spring 2015

Extended scaffolding: A more general theory of scaffolded cognition

Zachary R. Murphy
Purdue University

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Zachary R. Murphy

Entitled
EXTENDED SCAFFOLDING: A MORE GENERAL THEORY OF SCAFFOLDED COGNITION

For the degree of Master of Arts

Is approved by the final examining committee:

Daniel R. Kelly
Chair

Daniel Smith

Taylor Davis

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Daniel R. Kelly

Approved by: Matthias Steup 04/13/2015

Head of the Departmental Graduate Program Date

EXTENDED SCAFFOLDING:

A MORE GENERAL THEORY OF SCAFFOLDED COGNITION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Zachary R. Murphy

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Arts

May 2015

Purdue University

West Lafayette, Indiana

ACKNOWLEDGMENTS

I would like to express my wholehearted gratitude to my committee members. Without their thoughtful guidance at every step of the project, this thesis would not have been possible.

This work began as a term paper for the Fall 2014 class "Evolution and Externalism," co-taught by Dan Kelly and Dan Smith. I am indebted not only to their teaching of the material, but also to their support in helping me to formulate and refine the initial ideas that have culminated in this thesis.

I am sincerely thankful for Taylor Davis' careful reading and comments on numerous drafts, which have helped to refine my arguments dramatically on several dimensions.

To Dan, Dan, and Taylor: thank you for your support, encouragement, and guidance.

TABLE OF CONTENTS

Page

LIST OF FIGURES v

ABSTRACT vi

1 Introduction 1

2 Where is My Mind? 4

  2.1 Neuroprosthetics and Metaphysics 4

  2.2 EM and SM 10

    2.2.1 Extended Mind Thesis (EM) 11

    2.2.2 Scaffolded Mind Thesis (SM) 15

    2.2.3 EM vs SM 26

3 Extended Scaffolded Mind Thesis 34

  3.1 ESM 34

  3.2 How ESM relates to EM and SM 45

    3.2.1 ESM vs. EM 45

    3.2.2 ESM vs. SM 46

4 Case Studies: Applying ESM 48

  4.1 Human Minds 48

    4.1.1 Modern-day Humans 48

    4.1.2 Past Minds 50

    4.1.3 Future Minds 54

    4.1.4 Synchronic Group Diversity 55

  4.2 Non-Human Minds 56

5 Conclusion 62

REFERENCES 65

LIST OF FIGURES

Figure		Page
2.1	Cochlear Implant	5
2.2	Berger et al. Device	7
2.3	Andersen et al. Concept Device	9
4.1	DeepQA architecture	58

ABSTRACT

Murphy, Zachary R. MA, Purdue University, May 2015. Extended Scaffolding: A More General Theory of Scaffolded Cognition. Major Professor: Daniel Kelly.

New and emerging technologies called neuroprostheses are challenging our ideas about where one's mind ends and the environment begins. Cochlear implants, which completely replace the functioning of the inner ear, are now a common treatment for deafness. Berger et al. (2012) developed a device that replaces long-term memory in rats Berger et al. (2012), while Hampson et al. (2013) created a brain-machine interface that converts a desire to move one's arm into the motor neuron impulses required to achieve that movement Hampson et al. (2013)—both offering promising treatments for dementia, Alzheimer's disease, and paralysis. Deep brain stimulation is now a common way of regulating neural activity to manage muscle tremors in patients with Parkinson's disease. These devices completely replace parts of human anatomy we would normally consider to be performing cognitive processing. Are these devices themselves part of the cognitive system or do they just facilitate it? Are they parts of our minds or just sophisticated tools? More philosophically, these are metaphysical questions about where the boundary between the mind and the environment lies. Ultimately, this work establishes a principled way to set down such a boundary by developing a methodology for modeling potentially cognitive processes

using graph theory and applying graph-theoretical analytics to rigorously delineate mind from environment.

1. INTRODUCTION

New and emerging technologies called neuroprostheses are challenging our ideas about where one's mind ends and the environment begins. Cochlear implants, which completely replace the functioning of the inner ear, are now a common treatment for deafness. Berger et al. developed a device that replaces long-term memory in rats (2012), while Hampson et al. created a brain-machine interface that converts a desire to move one's arm into the motor neuron impulses required to achieve that movement(2013)—both offering promising treatments for dementia, Alzheimer's disease, and paralysis. Deep brain stimulation is now a common way of regulating neural activity to manage muscle tremors in patients with Parkinson's disease. These devices completely replace parts of human anatomy we would normally consider to be performing cognitive processing. Are these devices themselves part of the cognitive system or do they just facilitate it? Are they parts of our minds or just sophisticated tools? More philosophically, these are metaphysical questions about where the boundary between the mind and the environment lies. Ultimately, this work establishes a principled way to set down such a boundary.

I proceed in chapter 2 by first exploring the challenges these new technologies pose to our ideas about the mind and discuss two solutions concerning where to set the mind–environment boundary: those of Andy Clark and Kim Sterelny. Clark argues that cognition extends beyond the brain and skull to include anything, biological or not, that complements the cognitive processes occurring inside the skull. On the other hand, Sterelny argues that rather than saying extracranial processes are cognitive, we should say instead that they are just 'scaffoldings' for cognition: clever ways in which we organize our environment to ease processing demands on (intracranial) cognitive processing. Each justifies his position by an appeal to explanatory power. Clark thinks that nothing but his position gives adequate importance to the environment in accounting for cognition, while Sterelny argues that his theory not only gives adequate importance to the environment but accounts for more of the environment than Clark's theory.

In chapter 3, I put forward a new way of setting down a mind–environment boundary utilizing elements of Sterelny's scaffolding model. Sterelny himself does not take establishing a robust mind–environment boundary to be a primary goal. So while Sterelny makes some claims about the boundary of cognition, he does not offer a robust account of where this boundary lies. That said, Sterelny seems to limit cognition at the skull. I begin this chapter by showing how the scaffolding model itself is independent of Sterelny's claims about where the boundary is. I then put forward a more general scaffolding model—one that focuses not on agent–process scaffolding, but scaffolding as a relation holding between two processes. With this conception of scaffolding, we can model scaffolding using graph theory. Using graph-theoretical analytics, we can identify some core set of processes as the mind and all other processes as the environment. I take this use of graph theoretical modeling and analytics, together with a few assumptions to make the model align with current cognitive scientific understanding, to be a rigorous method of setting down the mind–environment boundary. Finally, this chapter finishes by showing how Extended Scaffolded Mind Thesis (ESM) is distinct from, and has explanatory advantages over, both of Clark's and Sterelny's models.

Chapter 4 includes a series of case studies, showing how ESM can be applied in several cases, including both human and non-human examples. In particular, I show how ESM can be applied to past, modern, and future humans. In doing so, I show how ESM adds to the explanatory power and rigor of the evolutionary theory of past, present, and future hominid cognitive development. Moreover, I demonstrate how ESM can be used to set mind–environment boundaries in radically non-human minds by applying ESM to IBM's Watson supercomputer. In doing so, I further illustrate the explanatory power and scope of ESM.

Finally, chapter 5 concludes. Ultimately, ESM is a principle by which to set down a mind–boundary distinction in a way that has more explanatory power than either Clark's or Sterelny's positions.

2. WHERE IS MY MIND?

2.1 Neuroprosthetics and Metaphysics

In recent years, new and developing technologies, such as cognitive prosthetics, have been challenging our ideas about where the mind ends and the environment or tools begin.

Cochlear implants can effectively replace the biological cochlea (Zeng, 2004). A healthy cochlea is lined with approximately 3,000 hair cells that are tuned to different frequencies. Different sounds resonate in different hairs, allowing the cochlea to register different sounds and convert them to neural impulses. Nerves from each of the hair cells coalesce into the auditory nerve, transmitting the encoded sound to the brain (most immediately to the temporal lobe). Cochlear implant candidates have deafness resulting from having significantly fewer hair cells, diminishing the neural impulse sent along the auditory nerve. In a cochlear implant, a microphone picks up sound waves, which are then digitized by a speech processor. This digital signal is then transmitted via radio waves across the skull to the implant, which sends the signal through an electrode fused to the auditory nerve (see fig. 2.1). In this way, a cochlear implant completely replaces the function of the cochlea, namely, converting sound waves into processed neural impulses.

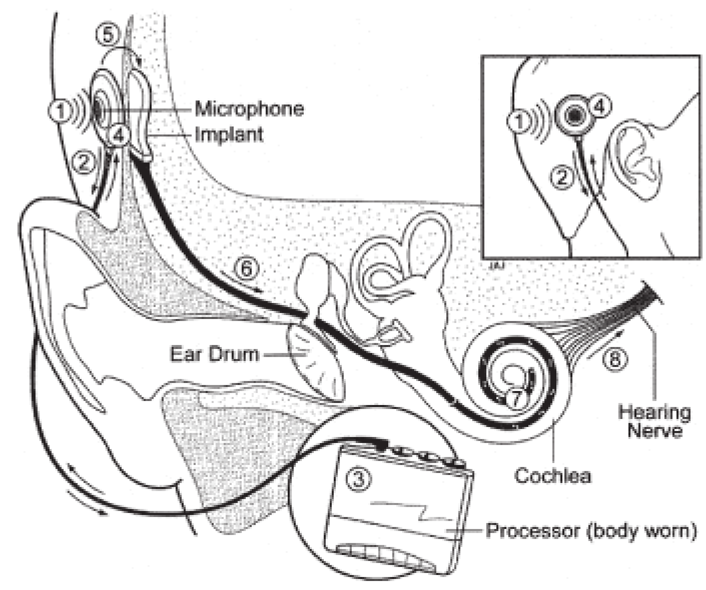

Fig. 2.1. Cochlear implant. Reproduced from Zeng (2004, p. 7).

Damage to the hippocampus can result in anterograde amnesia (the ability to form long-term memories), and is a common finding in patients with dementia and Alzheimer's disease (Berger et al., 2012, p. 198). Berger et al. developed a device that restored the ability to form long-term memories (2012). They found that neural activity recorded in the CA3 subfield of the hippocampus is highly predictive of memory responses in rabbits. They then disabled hippocampal output using MK801 (Dizocilpine), which left activity in CA3 unaffected but prevented the animal from forming long-term memories. They then re-introduced the electrical signal recorded from CA3 to the output neurons in CA1, resulting in the restoration of long-term memory formation abilities (Berger et al., 2012, p. 204-5). See fig. 2.2. A companion study showed that the same device applied to animals with intact long-term memory formation abilities results in increased performance, above normal biological capacities (Hampson et al., 2010). Hampson et al. applied this prosthesis to rhesus monkeys with the same methodology and found similar results (2013). This device effectively replaces or augments the CA1/CA3 subfields of the hippocampus, and has the potential to restore memory capacities to patients with dementia and Alzheimer's disease.

Deep brain stimulation of the subthalamic nucleus is now a common technique for regulating neural activity and managing tremors in patients with Parkinson's disease (Zibetti et al., 2011). An electrode is inserted into the subthalamic nucleus, the origin of muscle tremors in some patients with Parkinson's disease, and then connected to a (usually) subclavian pulse generator (Zibetti et al., 2011, p. 2328).

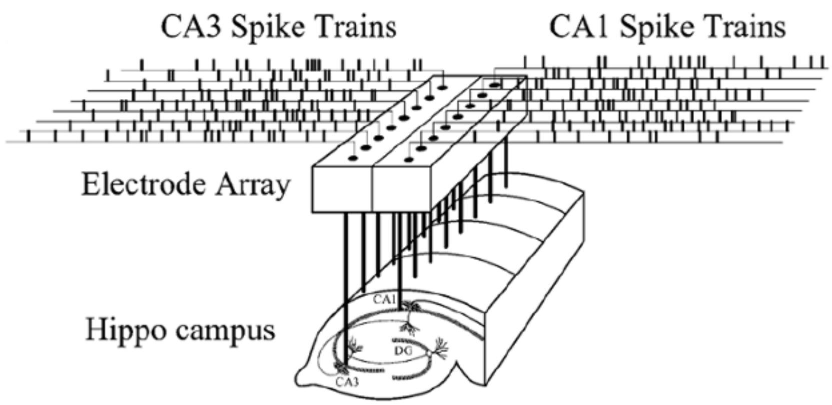

Fig. 2.2. Berger et al. device. Reproduced from Berger et al. (2012, p. 200).

The electrical impulses effectively negate the activity of the subthalamic nucleus, eliminating muscle tremors.

Andersen et al. have developed a model for a human-implantable device that takes over processing in the premotor cortex, which processes the desire to move into specific motor neuron impulses, and which has the potential to effectively cure some kinds of paralysis and create fully integrated limb prostheses. Most current neural limb prostheses are guided by motor signals, such as eye or facial muscle movement (Andersen et al., 2010, p. 171). While effective to a degree, these prostheses lag far behind natural limb movement. Faster, more integrated prostheses require direct access to neural activity further up the causal chain. Purposeful motor impulses are issued from the M1 motor cortex. However, such neuronal activity is largely devoid of goal information, since it consists of motor neuron commands. That said, the premotor cortex, particularly the posterior parietal cortex (PPC) and the parietal reach region (PRR), are highly predictive of goal-directed movement. The concept device developed by Andersen et al. registers this goal-directed activity in the premotor cortex and processes this activity into motor commands that can then be fed to efferent motor neurons or artificial limb prostheses (see fig. 2.3). Ultimately, the device replaces the functioning of the motor cortex.

Are these devices literally parts of our minds, or are they just tools our minds use to support their activities? This is ultimately a metaphysical question: what is the extension of the mind? Where is the boundary between mind and environment? One way to solve this question is to let our metaphysical commitments on the boundary of

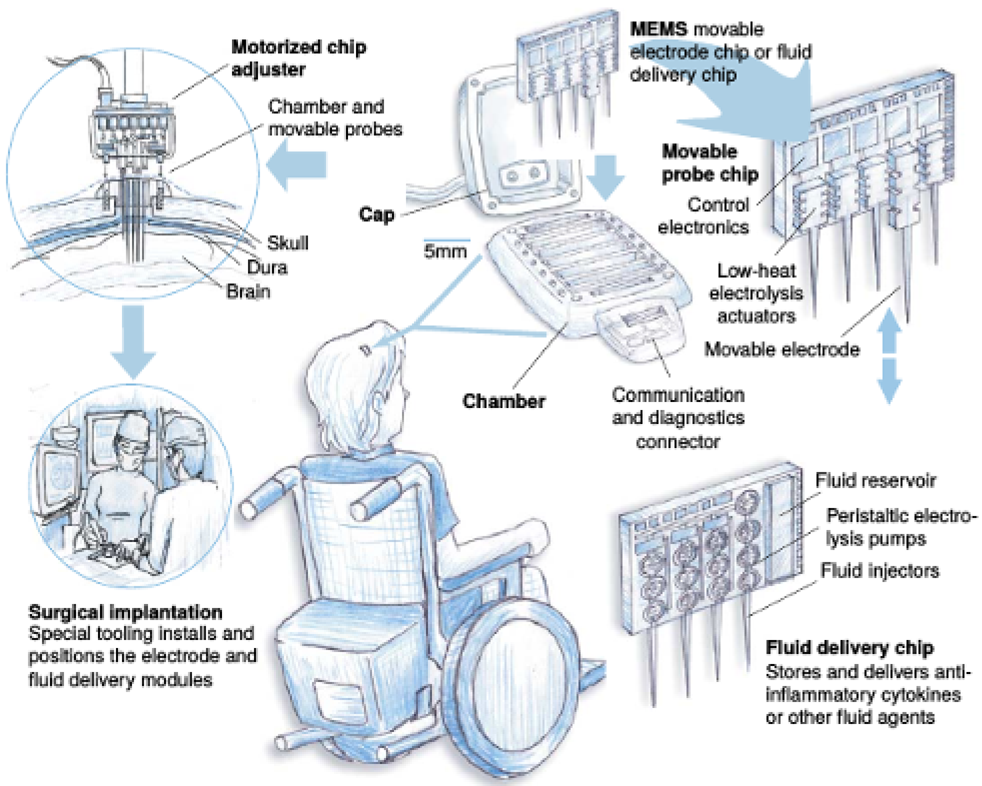

Fig. 2.3. Andersen et al. concept device. Reproduced from Andersen et al. (2004, p. 490).

the mind be led by what gives us the most explanatory power. That is, if saying that a neuroprosthesis is part of the mind yields greater explanatory power than saying it merely facilitates cognition, then we should consider the neuroprosthesis to be part of the mind. Similarly, if saying that a neuroprosthesis is merely an aid to cognition rather than a part of cognition itself yields greater explanatory power, then we should not call the neuroprosthesis itself cognition. Ultimately, this work presents a way to draw the boundary between mind and environment along this line of thought. Before stating this answer, though, we must explore previous answers from which it is built.

2.2 EM and SM

At least two camps of thought have emerged on how to set down the mind–environment boundary so as to maximize explanatory power. Andy Clark's Extended Mind Thesis claims that the mind can, and does, extend beyond the brain. On the other hand, Kim Sterelny's Scaffolded Mind Thesis says that cognition stays bounded by the skull, and the things EM points to as cognition are really just clever ways of facilitating cognition. The remainder of this chapter first looks at each of these theses in detail, then recounts the exchange of blows over which one has greater explanatory power.

2.2.1 Extended Mind Thesis (EM)

The Extended Mind Thesis claims that cognitive processes can be instantiated outside of the skull. There have been at least two prominent formulations of EM: the parity and complementariness formulations, which I explain below.

Parity formulation

The Parity Formulation of EM (EM_P) was initially proposed by Andy Clark and David Chalmers in "The extended mind" (Clark & Chalmers, 1998), and was one of the first theses to get EM off the ground. To motivate their claim, they consider the cases of Inga and Otto. Suppose Inga wants to go to the Museum of Modern Art (MOMA). She consults her memory for a second, recalls the location of the museum, and takes off (Clark & Chalmers, 1998, p. 12). Presumably, the mechanism realizing Inga's memory is some set of neurons. In this case, we seem to have no trouble saying that these neurons are, in some relevant sense, part of her mind and performing cognitive processes; her belief that the MOMA was at such and such a place seems contained in these memory neurons.

Now consider Otto, who cannot retain long-term memories due to having Alzheimer's disease. To compensate for this deficit, Otto carries around a notebook in which he writes down things he wants to remember later (Clark & Chalmers, 1998, p. 12). Let's assume written in his notebook is the location of the MOMA. Suppose Otto wants to go to the MOMA. He consults his notebook quickly, recalls the location of the museum, and takes off (Clark & Chalmers, 1998, p. 12). Clark and Chalmers point out that Otto's notebook serves the same role as Inga's memory neurons: containing the information of where the MOMA is (Clark & Chalmers, 1998, p. 15). Moreover, given how tightly Otto is integrated with his notebook, the *only* possibly relevant difference between Otto and Inga is that Otto's memories are stored in a notebook while Inga's memories are stored in the brain. However, such a difference seems superficial: we should, Clark and Chalmers claim, say that Otto's notebook is just as much a part of his mind as Inga's memory neurons are—we should say Otto's notebook is, in a literal and robust way, part of his mind (Clark & Chalmers, 1998, p. 13). Generalizing, they say that we should not be brain chauvinists: "there is nothing sacred about skull and skin" (Clark & Chalmers, 1998, p. 14). To say that Inga's memory neurons are part of her mind while Otto's notebook is not makes an arbitrary, and hence unnecessary and unjustified, boundary of cognition at the skull.

More formally, they state what has become known as the Parity Principle (PP):

> If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* part of the cognitive process. (Clark & Chalmers, 1998, p. 8)

That is, anything outside the skull which functions just like something within the skull should be considered to be literally part of the mind. Here we can clearly see the Functionalist underpinnings of EM_{P} (Clark, 2008a, p. 46).

Several problems have been raised against EM_P, of which I'll mention just one of the most pressing. PP only allows for cognitive mirroring: the only sorts of mental processes that can exist beyond the skull are those that are already inside the skull. Strictly adhering to PP, any difference in functioning between intracranial and extracranial processes is enough to say that the extracranial process is *not* cognitive. The only kinds of cognitive processes that are possible are those that are already instantiated intracranially. If there is no intracranial functional equivalent of an extracranial process, then the extracranial process is automatically disqualified from being cognitive (Sutton, 2010, p. 196). However, this seems contrary to the line of argument Clark wants to pursue, since he wants to move away from a commitment to the priority of intracranial processes. So, EM_P seems ill-suited to serve as a general EM principle of what processes should be included as part of the mind.

Complementariness formulation

The Complementariness Formulation of EM (EM_C) arose as a response to problems with EM_P, and has been championed (for my purposes) by John Sutton (Kirchhoff, 2012, p. 290). Sutton presents his 'Complementarity Principle:'

> In extended cognitive systems, external states and processes need not mimic or replicate the formats, dynamics, or functions of inner states and processes. Rather, different components of the overall (enduring or temporary) system can play quite different roles and have different properties

> while coupling in collective and complementary contributions to flexible thinking and acting. (Sutton, 2010, p. 194)

This is to say that so long as a process is integrated into the cognitive system in the right way—whether it resembles some mental process internal to the skull or not—it should be considered a cognitive process. This line of thought focuses on saying that what is important in deciding whether or not some process is part of some cognitive system is not whether the process resembles anything that is already going on within that system, but instead focuses on whether or not that process is in the right functional relationship to the cognitive system.

So, for example, to know whether or not Otto's notebook is literally part of his mind, we shouldn't look for some isomorphism between the notebook and some undisputedly cognitive process within Otto's skull. Instead, we should look at how the notebook interacts with Otto's intracranial cognitive processes. It is highly coupled with Otto's intracranial cognitive system, he has access to it at any time, and he is highly dependent on it. These factors are quintessentially complementary to the intracranial cognitive processes, and so—according to EM_C—we should consider Otto's notebook to be part of his mind.

We can view EM_C as a refinement and subsuming of EM_P (Sutton, 2010, p. 206) As Sutton et al. say, "Complementarity best captures the spirit of extended cognition...The human brain is a leaky associative engine (Clark 1993), shaped in both evolution and development so as actively to integrate and coopt external resources such as media, objects, and other people" (2010, 525-6). Any process that meets the

PP conditions to be considered a cognitive process is, de facto, a process that complements intracranial cognitive processing and so also satisfies the conditions of EM_C. Moreover, EM_C is explicitly not committed to the cognitive mirroring requirement of EM_P, since what is important is the functional integration of an extracranial process rather than what function the process itself is performing.

Recently Clark has advocated a view closer to EM_C than EM_P, and has even claimed that EM_C was his intended thesis all along (Clark, 2008b, 114; Clark, 2010, 52). He says that PP was meant, from the beginning, as a "'veil of ignorance' style test meant to help avoid biochauvinistic prejudice" (Clark, 2008b, p. 77). This seems plausible, given the presentation of several 'coupling criteria' in Clark & Chalmers (1998), such as portability and reliability, which suggest a more nuanced theory than the EM_P formulation of EM.

Given these considerations, the target formulation of EM I will use for the rest of this work will be EM_C, and I use EM and EM_C interchangeably unless stated otherwise.

2.2.2 Scaffolded Mind Thesis (SM)

The Scaffolded Mind Thesis thesis, put forward by Kim Sterelny, claims that cognitive processes are not instantiated outside of the skull, and the extracranial processes that EM points to are nothing more than clever ways of organizing one's environment to facilitate and ease real cognitive processing (occurring solely within the skull). In "Minds: extended or scaffolded?" (2010) Sterelny makes this response to EM—in particular to EM_P-type theses, although much of what he says carries over to EM_C-type theses. He sets out to argue that

> the canonical extended mind cases are continuous with other cases, cases in which there is environmental support of cognition, but which are not plausibly treated as constituents of agents' minds. Moreover, the dependence of cognitive competence on extra-somatic resources turns out to be a special case of a more general phenomenon [namely scaffolding]. (Sterelny, 2010, p. 466)

Scaffolded digestion

Thinking about cognitive extension is messy. To ease into an explanation of cognitive scaffolding, Sterelny takes a helpful detour into digestive scaffolding first, and I will follow him on this explanatory path. Consider human digestion and how it has changed over time. Digestion requires the breakdown of foodstuffs into components to be absorbed into the bloodstream. We have processes internal to the body that perform this function: teeth begin a brute, manual breakdown of food into smaller pieces. Saliva contains amylase, which starts to break down starches. Stomach acid is perhaps the most vivid process of the chemical breakdown of foodstuffs. The small intestine has a specialized lining that allows foodstuffs to enter the bloodstream, and the colon recovers fluids from the digestive matter. Every mechanism in this process is evolutionarily optimized: the size and shape of teeth, chemical constitution of saliva, size of stomach, strength of stomach acid, and the length of the intestines are all optimized for the type of foodstuffs that enter the body. Thus, humans and cows have different digestive mechanisms due to differences in what goes in the mouth. Herbivores, carnivores, and omnivores all have different types of hardware, each type optimized to what goes in the mouth.

However, there is something that further sets humans apart from herbivores and carnivores, and even other omnivores: humans cook their food. We prepare food in a wide range of ways: cutting, pounding, soaking, grinding, fermenting, baking, boiling, roasting, pickling, and so on (Sterelny, 2010, p. 467). Well-prepared food is certainly a hedonistic pleasure, but there is more to be said for its importance. By preparing food, humans have changed the norm of what enters their mouths: instead of eating raw foods, most foods humans eat are prepared in some way. What is important here is that these ways of preparing food are ways of breaking down the foodstuffs to be more readily digestible—preparing food performs processes of the digestive system. Cutting performs the processes done by chewing; baking, roasting, and boiling perform the processes done by digestive juices; and so on. So, processes that the human digestive system would otherwise have to perform are instead offloaded into the environment via things like fire, ovens, knives, boiling water, etc.

Moreover, as we have evolved this offloading of digestive processing onto the environment has changed the human body. Since we cut our food into small chunks, we don't need massive, powerful teeth. Further, massive and powerful teeth become an evolutionary burden: developing such teeth takes resources that could have been expended elsewhere. Similarly, our "under-powered jaws, short gut, small teeth and mouth" are the result of our gustatory niche—one that depends on food preparation (Sterelny, 2010, p. 468). Our reliance on food processing can be readily seen in the difficulties raw foodists have in maintaining healthy levels of food intake and nutrient absorption (Sterelny, 2010, p. 467). So, in offloading digestive processing into the environment, we have become dependent on these extrasomatic processes.

One thing we *could* say about our food preparation is that since it performs the same processes as body-based digestion, food preparation *is* digestion. This would make cooking a case of 'extended digestion.' Sterelny points out, however, that he knows of no one who is prepared to make such a statement (Sterelny, 2010, p. 468). Instead, Sterelny says that "[o]ur digestive system is *environmentally scaffolded*" (Sterelny, 2010, p. 468, emphasis in original). To say that food preparation is scaffolding for digestion and is not digestion itself is to say that food preparation is a clever way in which humans have organized and affected their environment in such a way as to ease processing demands on the actual digestive system itself.

Scaffolded cognition

Sterelny's claim concerning extended cognition is that the cases that EM points to are to the mind as food preparation is to digestion. They are cases of offloading processing into the environment, but merely as scaffolds for cognition and not as parts of the cognitive system themselves. As he says, "The scaffolded mind hypothesis proposes that human cognitive capacities both depend on and have been transformed by environmental resources" (Sterelny, 2010, p. 472). So, SM says that the environment is fundamentally important for cognition, but not in the way EM holds.

Here Sterelny tells the story he has been advocating in the decade leading up to "Minds: extended or scaffolded," which shows how "intergenerational social learning profoundly shapes our minds and lives" (Sterelny, 2010, p. 470). This story is fleshed out in his 2003 book *Thought in a Hostile World*, in which he adopts the niche construction model developed by Kevin Laland, John Odling-Smee, and Marcus Feldman. A fundamental concept of evolutionary theory is that animals adapt to their environments and form niches, i.e. particular ways of interacting with the environment in order to meet biological needs. As an organism's niche changes, the organism adapts to those changes. Agents of niche change might include climate change, for example, or the introduction or elimination of another species from an environment. However, some organisms engage in 'niche construction,' wherein they themselves act as agents of change for their own niches. Typical products of niche construction are burrows, nests, and webs, which all "modify the impact of the environment on their builders" (Sterelny, 2003, p. 147). So, for organisms that engage in niche construction, there is a two-way adaptive process between the organism and its environment. The organism adapts to its niche, but then changes its niche in some way. It then adapts to this new niche, changes that niche again, adapts to this new niche, and so on.

Sterelny's claim is that humans engage in cognitive niche construction: "human cognitive capacities both depend on and have been transformed by environmental resources" (Sterelny, 2010, p. 472). Over evolutionary time, humans have developed ways to organize their environment in ways that offload some cognitive processing from intracranial architecture onto environmental entities. He calls these environmental supports 'cognitive scaffolding.' For example, long-term memory is certainly an important component of human life. However, its biological hardware (neurons) is metabolically expensive. To have the benefits of long-term memory yet not expend metabolic resources on biological hardware, we store information in the environment that is readily accessible when the need for it arises. One way we do this is through writing. Writing stores information in the environment for access later, but does not take up expensive neural resources. Moreover, through writing a single individual can 'recall' more information than he/she could ever learn and memorize. So, not only does writing ease cognitive processing demands, but it also facilitates access to more information than the human mind can store at once. So, notes, books, etc. are cognitive scaffoldings since they are important for and support cognition, but are not themselves cognitive processes. In this way, humans have engaged in cognitive niche construction through the development of cognitive scaffoldings.

Sterelny suggests that there are at least three dimensions of cognitive scaffolding: trust, individualization, and entrenchment. Each of these is a graded functional relationship between the agent—i.e. mind—and some part of the environment (Sterelny, 2010, p. 473). That is, being trusted, individualized, entrenched, etc., is a matter of degree rather than being a bivalent property.

Trust Sterelny defines the trust dimension of scaffolding as "the agent's assessment of the reliability of their access to a resource and the reliability of the resource itself" (Sterelny, 2010, p. 474).[1] That is to say, the more an agent believes that a cognitive resource[2] is always at hand and contains reliable information or processing, the more that resource functions as a scaffold for the mind. For example, Otto puts a high credence in the assumptions that (1) his notebook will always be there when he needs it, and (2) the information in his notebook is reliable. So, Otto's notebook falls fairly high on the trust dimension. On the other hand, my notebook is not always with me, and I think it is not unreasonable to think someone could take it and modify the information in it. So, my notebook has less trust than Otto's, and is in that sense less of a scaffolding. As this makes apparent, trust is on a gradient: resources can be more or less trusted, and this correlates with their being more or less scaffolded.

Individualization A cognitive resource is individualized when it is utilized by one cognitive system, i.e. one person. Often resources are customized to work for a single person. Sterelny talks about professional chef's knives which are custom-molded to their hands (Sterelny, 2010, p. 475). Such knives only work for the person they were designed for; nobody else can pick them up and work with them fluently. Similarly, Otto's notebook is of use only to him. While the notes jotted down may be meaningful for him, it may not be useful for just anyone who comes along and flips through

[1] It seems like there might really be two dimensions here: 1) trust in reliable access to the resource and 2) trust in reliable information contained in the resource. We could very well treat these as two different dimensions—Sterelny is not committed to these exact dimensions or the number of dimensions.

[2] Terminological point: I use 'cognitive resource' to pick out a process that is not itself cognitive but is nonetheless important for cognition.

the notebook. On the other hand, anyone could pick up my smartphone, find the calendar, and gain the same information about my schedule that I get when I look at it. So, other people would have no trouble using my calendar, and thus it is less individualized than Otto's notebook. Again we can see that individualization is on a continuum.[3]

Entrenchment A cognitive resource is entrenched when a cognitive system is highly attuned to it, and were it replaced with a similar but not exact duplicate the cognitive system would have trouble (or at least more difficulty) in using it. As Sterelny says, again using the example of professional chefs,

> Modification will often be mutual. The cook's physical routines and skills adjust to the weight, balance and sharpness of her knives; the cook's movements are adapted to her knife, just as the knife has been individualized to her prior preferences. If forced to use different knives, an expert will cope, but physical routines will be less comfortable and precise." (Sterelny, 2010, p. 475)

The more an agent is attuned to a specific resource where another, albeit similar resource simply will not do, that resource is entrenched. Again, entrenchment is on a gradient, depending on how well the agent could cope with a different but similar

[3] Sterelny is not clear on why he thinks this is an important dimensions, but something akin to it runs through Clark's works as well. So, arguing that individualization is not important would tell just as much against Clark as Sterelny. Moreover, Sterelny is not committed to any specific set of scaffolding dimensions and the inclusion of individualization among them. In Sterelny (2004), he may even be going the other way, saying "Clark underplays the importance of non-exclusive use of epistemic artefacts. Many of our most important cognitive tools are common-use tools, not parts of coupled systems" (Sterelny, 2004, p. 245).

resource. So, suppose we give Otto a different notebook with all the same information in it, but organized in a different order. Presumably, he would have more difficulty using such a notebook than if he had his own. He would likely ask for his original notebook back. So, Otto's notebook has a good degree of entrenchment.

As we can see, Otto and Otto-type cases fit easily into Sterelny's scaffolding space. So, Sterelny says that Otto-type cases exist at the maximal point of the 3D scaffolding space (Sterelny, 2010, p. 480). By placing Otto-type cases here, Sterelny makes it clear how such cases are merely a special case of a more general phenomenon; they rank high on all three of the scaffolding dimensions, but this is all that separates them from cases that lie elsewhere in the scaffolding space.

Sterelny is careful to say that there are *at least* these three dimensions to scaffolding—there could be more. To have a neat way of talking about a process being more or less scaffolded along this multitude of possible dimensions, I use 'level of integration' as a catch-all term for the degree to which a device or process functions as a scaffold for a cognitive system, or the degree to which some external resource has been integrated into the workings of the mind, independently of what the specific dimensions of scaffolding may be. So, a 'highly integrated' process is one which ranks high on the dimensions of scaffolding, whatever they happen to be in a fully fleshed out description of what constitutes cognitive scaffolding.

Sterelny uses this scaffolding model in service of accounting for hominid cognitive development. Here I show an example of how Sterelny applies SM in service of his explanatory goals. Although it is not Sterelny's direct goal, I will ultimately show

how my theory adds an interesting layer of explanation to Sterelny's explanation by tracing the expansion of the boundary of cognition throughout each of the cognitive histories Sterelny offers.

In a chapter entitled "Heterogeneous Environments and Variable Response" Sterelny (2003), Sterelny sets out to offer an explanation of human developmental plasticity: the exceptionally high degree to which differences in environment during the developmental stage of hominids (presumably the time up to and including adolescence) can impact cognitive functioning. That is, human cognition is widely variable depending on what environment the individual is raised in, to a degree unparalleled by any other organism, and this anomalous property of hominids needs to be explained.

He pushes back against Cosmides and Tooby's claims that we have cognitive architectures adapted specifically to the Pleistocene, and instead argues that developmental plasticity is itself an adaptation to environmental instability:

> [T]here is good reason to believe that cognitive and neural developmental plasticity is adaptive; that human life-history considerations indicate that we have evolved to extend our developmental period; that there is paleobiological evidence that hominids experienced increasing climactic and ecological instability. We do not just happen to be part of a highly plastic lineage; our developmental plasticity is an adaptation. (Sterelny, 2003, pp. 170-171)

That is, in response to increasing environmental variability in the Pleistocene era, hominids developed a lengthened developmental period in which to adapt to any environment in which the individual was born. Instead of adapting to Pleistocene conditions in particular (as Cosmides and Tooby argue), hominids instead developed ways to adapt to a wide array of environments, whether Pleistocene or modern. He points to three dimensions of this plasticity: automated skills, affect, and neural plasticity (Sterelny, 2003, p. 163). Automated skills are developed during childhood, which then become permanent in adulthood. Similarly, aversions and disgust responses (affects) are acquired in childhood and are generally maintained throughout adulthood. Moreover, human brains are themselves plastic in that neuronal patterns are formed during adolescence and generally remain until death. As Sterelny notes, although we know that human brains are plastic, the degree to which humans brains are plastic is still an "open empirical question" and there is much more research to be done (Sterelny, 2003, p. 164). Nonetheless, hominid brains are malleable to some significant degree through adolescence, and this malleability allows hominids to take on skills, affects, and other behavioral and physiological responses that persist throughout the individual's lifetime.

Importantly, for Sterelny, this developmental plasticity is molded not only by the environment in which an individual happens to be born, but also by the products of cumulative downstream niche construction:

> Human brains are developmentally plastic, so transforming hominid developmental environments transformed hominid brains themselves. As

> hominids remade their own world, they indirectly remade themselves. (Sterelny, 2003, p. 173)

That is, hominids began to take advantage of the exaggerated developmental plasticity to modify hominid cognitive functioning through intergenerational learning. For humans, the neural patterns set during one individual's period of developmental plasticity do not just stay with the individual. These patterns are subsequently passed down to later generations during their developmental periods through intergenerational learning. This constitutes a special form of niche construction, where members of one generation actively influence the cognitive architecture of the next generation. In this way, hominids have ha a direct impact on the development of their own minds and have effectively "remade themselves" (Sterelny, 2003, p. 173).

2.2.3 EM vs SM

Clark (2008b) and Sterelny (2010) contain an exchange between Clark and Sterelny over which of EM and SM has more 'heuristic value.' Sterelny appears to be the first to bring in this terminology of 'heuristic value,' although he does not make clear what he intends by it (neither does Clark, for that matter). As I will show, I think we can interpret Clark and Sterelny as being in agreement about the fact that what is at issue is whether EM or SM has more explanatory power. So, I take them to each be positing reasons why their respective theory has more explanatory power than the other, in virtue of where it draws the mind–environment boundary.

Clark's Arguments

In Clark (2008b), Clark argues that the EM framework is more compelling than any of the other several theories that have been put forward recently, including Sterelny's SM. The basis of his arguments is that no theory other than EM gives adequate importance to the environment (Sterelny, 2010, p. 479). He states:

> What [EM] allows us to see clearly is that where ongoing human cognitive activity is concerned, there are usually *many* boundaries in play, *many different kinds* of capacity and resource in action, and a complex and somewhat anarchic flux of recruitment, retrieval, and processing defined across these shifting, heterogeneous, multifaceted wholes. To identify the bounds of cognition with the bounds of the brain/CNS, or even with those of the biological organism [as in SM], is to elevate just one or two of these many boundaries and interfaces to permanent cognitive glory at the expense of all the rest. (Clark, 2008b, p. 138, emphasis in original).

Here Clark makes the claim that anything that limits cognition at the skull discounts the importance of extracranial resources for cognition. Although it is not language he explicitly uses often, I think it is reasonable to say that he is making a claim about explanatory power: he sees extracranial resources as having a fundamentally important role in cognition, and all theories except EM fail to explain this importance.

In particular, he argues that Sterelny's SM fails to adequately account for the important role extracranial processes play for human cognition. In response to Sterelny's account of downstream niche construction through intergenerational teaching, Clark states:

> Sterelny's emphasis is thus very much upon the direct neural consequences of the culturally and artifactually scaffolded training regimes applied to young human minds. But although such consequences are surely of the utmost importance, they do not yet exhaust the cognition-transforming effects of material artifacts and culture. For many of the new cognitive regimes supported by our best bouts of incremental epistemic engineering seem to resist full internalization. It is no use, as Ed Hutchins (personal communication) points out, trying to imagine a slide rule when you need to work out a log or cosine! (Clark, 2008b, p. 68)

The idea here, I gather, is that accounting for how extracranial resources are developed, passed down, and refined still does not account for why we are dependent on the material artifact itself to perform a task. That is, according to Clark, Sterelny has given us a story about how slide rules developed and how a knowledge of slide rules is passed down from one generation to the next and improved upon, but not a story about why we cannot perform some mathematical calculations without having a slide rule literally in hand. However, according to Clark, EM is perfectly capable of explaining this: not having a slide rule is just like having part of your brain missing—the hardware carrying out the processing is not present. So, EM better explains the importance of extracranial cognitive resources than SM.

Sterelny's Arguments

In Sterelny (2010), Sterelny takes up Clark's argument and claims to show that SM need not underplay the importance of the environment; he thinks that scaffolding does give adequate importance to extracranial processes. SM is committed to the claim that cognitive scaffoldings, although not themselves cognitive processes, are necessary for certain cognitive processes to occur. Moreover, when an organism constructs a niche, it becomes functionally dependent upon those environmental modifications.

For example, SM can account for why we cannot calculate logs and cosines without a material slide rule. SM holds that slide rules developed as a way to store a wealth of unit conversion information in the environment when storing all of this information in on-board long-term memory or performing the complex calculations in the head are both expensive. So, the reason we cannot perform log or cosine calculations without a material slide rule present is essentially the same answer given by EM: the information is in the slide rule, and the slide rule isn't here at the moment. The difference between SM and EM in this case is that whereas EM claims that the slide rule constitutes a cognitive process (in particular long-term memory), SM claims that the slide rule is just an aid to cognition. According to SM, calculating a log or cosine with a slide rule is not, strictly speaking, performing a calculation—it is a perceptual problem wherein one looks at the information contained on the slide rule and accepts that piece of information as correct. In this way, Sterelny argues that SM does not understate the importance of slide rules to calculating a log or cosine, and in particular it does explain our dependence on actual, material slide rules.[4] Clark is right to say that to explain how slide rules developed is not exactly to explain why people depend upon them. However, to explain how individuals develop in environments containing slide rules *is* is to explain eactly why people depend upon them. So, Sterelny defeats Clark's claim that only EM can give adequate importance to the environment.

Sterelny then turns to arguing that SM has explanatory power EM lacks. He begins by saying that while there is nothing obviously wrong with calling the maximal point of the scaffolding space—occupied by Otto-type cases—literally part of the mind, "[i]t obscures the fact that extended mind cases are special cases of a general phenomenon" (Sterelny, 2010, p. 480). His argument here seems to be that if we were to priviledge this maximal point of the scaffolding space (as EM does), this discounts the importance of the rest of the range of cognitive scaffolding. Focusing on just maximal cases leads us to think of less integrated cognitive resources as less important. But, as Sterelny has shown, these less-than-maximal scaffoldings are nevertheless sometimes vitally important, and often serve as the early forms of what eventually develop or evolve into highly integrated scaffoldings.

[4]Similar arguments can be made for calculators as a way of calculating logs and cosines. One difference, though, is that whereas a slide rule simply stores information, a calculator actually performs symbol manipulations, which can then be read as calculations. So, while slide rules offload long-term memory processing into the environment, calculators offload calculation (i.e. some form of abstract symbol processing) into the environment. Beyond this, though, the cases are parallel: calculators reduce calculation to a perceptual problem just like slide rules (Sterelny, 2007).

Moreover, not only does EM's privileging of Otto-type cases distract from the importance of less-than-maximal scaffoldings ('partial integrations'), but EM has nothing to say about such cases. EM does not have the resources—short of adopting a scaffolding framework—to talk about anything less integrated than Otto cases as anything other than relatively uninteresting parts of the environment. That is, EM accounts for things like Otto's notebook by saying that they are parts of the cognitive system. This works when the resource under consideration is at the maximal scaffolding point. However, EM does not seem to have anything to say about something not maximally scaffolded, like my notebook. My notebook is certainly important for my cognition: it stores information so that I can access it later and without it I would be missing out on important pieces of information, but it is not integrated with my mind to the extent that Otto's notebook is integrated with his.

EM faces a dilemma here: it must either claim that my notebook is part of my mind, or claim that it is not part of my mind. If EM says that my notebook is part of my mind, the same reasoning can be used to say that an oven is part of the modern human's digestive system. If we take this claim to be metaphorical and mean only that ovens have an impact on the modern human's digestive system, then this is not so radical of a claim. Clark, however, intends these claims to be taken literally: if Clark is to be consistent, he must say that ovens are literally part of the human digestive system. However, to Sterelny and to me this seems obviously wrong (Sterelny, 2010, p. 468) One implication of this view that draws out why I find Clark's position here untenable is that it seems obviously wrong that gastrointestinal physicians, who are required to be well-educated on every aspect of human digestion, should be required to also understand oven function, construction, and maintenance in order to practice their specialty. On the other hand, if EM says that my notebook is not part of my mind, then it does not have any resources to explain why it is nonetheless important for my thinking. So, EM either leads to an abuse of the term 'cognitive' or fails to explain large classes of phenomena.

Sterelny, however, uses the entire scaffolding space and the claim that there is a continuum between my notebook and Otto's to explain this. My notebook is important for my cognition because I have stored information in it so that I do not have to spend expensive neural resources on remembering that information yet can still have access to it later. The only difference between my notebook and Otto's notebook is that his is more trusted, individualized, and entrenched—the difference is merely a matter of degree. So, according to Sterelny, SM is more "heuristically compelling" than EM since it has an account of the importance of less-than-maximal scaffoldings, which EM doesn't (Sterelny, 2010, p. 480).

Sterelny uses the terminology of 'heuristic advantage.' However, he does not elaborate on what he means by this and what sort of value he sees here. That said, I think it is reasonable to interpret him as making claims about explanatory power: he has shown that SM can explain and account for phenomena that EM can't.

Furthermore, Sterelny thinks that the EM picture cannot answer what he takes to be the interesting and important questions:

> Are these the only dimensions of importance? To what extent are they independent? What are the dynamics of movement in the space? Under what circumstances, for example, do collectively used resources become segmented into single-user resources and vice versa? How do resources become individualized and entrenched? How to actual levels of trust covary with trustworthiness? (Sterelny, 2010, p. 480)

EM has trouble answering these questions, since it does not have the tools to talk about partial integrations. On the other hand, SM is a better model for these questions since it “focuses on the space itself” (Sterelny, 2010, p. 480). So, SM has more explanatory power than EM since it has answers to questions that EM doesn’t.

3. EXTENDED SCAFFOLDED MIND THESIS

I think that SM is a useful model for setting down the mind–environment boundary, and my solution uses it as a foundation. However, I see Sterelny's notion of scaffolding as a special case of a more general notion of scaffolding. Using this more general notion of scaffolding, we can formulate a robust and principled way of setting down the mind–environment boundary. The first half of this chapter is an explanation of this more general notion of scaffolding and how we can go about using it to setting down mind–environment boundaries rigorously. The second half of this chapter shows how this theory is both distinct, and has explanatory benefits over, both SM and EM.

3.1 ESM

My goal is to use the scaffolding framework to set down a rigorous and principled mind–environment boundary through what I call the Extended Scaffolded Mind Thesis (ESM). Before starting on a statement of ESM, however, it is important to be clear on the work ESM is doing and, equally importantly, the work it is not doing. I am interested in setting down the mind–environment boundary, but not necessarily identifying minds in the first place. To that end, ESM does not tell us what beings are minded and which ones are not, but instead—given some being that we grant has a mind—it tells us where the boundaries of that being's mind are. So, instead of attempting to give necessary and sufficient conditions for a being to be minded, ESM begins with the assumption that an entity is minded and delineates that mind from its environment.

Sterelny uses the scaffolding framework as a relation between the agent and the environment. Throughout his work, Sterelny makes use of a distinction between agents and scaffolds. Presumably, his use of 'agent' maps onto this work's use of 'mind.' Every process outside of the agent is somewhere on the scaffolding space, ranging from minimal to maximal importance to the agent on whatever dimensions of scaffolding there happen to be. So, Sterelny's agent–scaffold distinction maps onto the mind–environment boundary used in this work. In this way, scaffolding for Sterelny is a relationship that takes the form 'process p scaffolds agent a.' What makes a process scaffolding for the agent is that the agent depends on that process or at the very least that process makes the agent's own processing easier. This is a relationship of degree, and there are multiple dimensions along which it varies, such as trust/reliability, individualization, and entrenchment. Call this 'agent–process' scaffolding.

However, there is a more general (and ultimately more useful) notion of scaffolding we can employ, which focuses on scaffolding as a relationship between two processes, rather than between an agent and a process. Ultimately, I'll show that we can use this notion to build back up to understanding scaffolding at the level of agent and process, but in a more principled and robust way than Sterelny offers. Scaffolding, for ESM, is the relationship holding between two processes p_1 and p_2 such that when p_1 scaffolds p_2, p_1 eases p_2's processing demands. Call this 'process–process scaffolding.' Process–process scaffolding is still a matter of degree, and there are multiple dimensions along which it varies. Many of these align with the dimensions Sterelny sets out, and it preserves the same flexibility regarding which dimensions are salient. We can say that p_1 is trusted/reliable for p_2 when p_2 generally does not verify input from p_1 against other inputs, and p_1 is generally available whenever p_2 needs its input. We can say that p_1 is entrenched for p_2 when p_2 cannot operate effectively without p_1. Individualization is tougher to see, but we can still define it as a sort of meta-scaffolding dimension. Say that p_1 is individualized for p_2 when p_1 ranks highly on other scaffolding dimensions for p_2 and only p_2. In this way, we can define dimensions of scaffolding in terms of a process–process relationship rather than an agent–process relationship.

A feature of process–process scaffolding is that it can be modeled on (and is perhaps easier to grasp through) graph theory. A graph is composed of nodes and (possibly multiple) connections between them. In this case, the nodes would represent processes and the connections (called 'edges') represent dimensions of scaffolding.[1] Moreover, the connections can be directed: we can say that a relation is one-way and nonsymmetrical, as the scaffolding relationship is. We can say that p_1 scaffolds p_2, but p_2 does not scaffold p_1. Even further, we can give different connections different weights or strengths. Obviously numerical weights would be arbitrary for scaffolding, but they can still be meaningful in relation to each other. A relevant question,

[1] Although not concerned so much with scaffolding as the flow of information, modern functional neuroanatomy operates on much the same model.

though, is whether the range of scaffolding weights is bounded on either end: is there a minimal or maximal level of scaffolding? My inclination is to say that there are no such boundary points: there is no maximal or minimal level of scaffolding, but the model I develop works equally well whether we say such limits exist or not. Furthermore, this process network is transitive, with the transitive weight being no greater than the weakest scaffold: if p_1 scaffolds p_2 with weight w_a and p_2 scaffolds p_3 with weight w_b, then p_1 scaffolds p_3 with weight $\min(w_1, w_2)$. All together, we can model the world as a multidimensional, directed, weighted network of processes, related to each other through scaffolding relationships along each of the scaffolding dimensions. Call this graph of all processes and their multi-dimensional scaffolding relationships the 'process network.'

Measuring scaffolding weight between two processes would take the form of empirical testing regarding the strength of the quality associated with each dimension of scaffolding. For instance, measuring the weight of entrenchment between two processes amounts to empirical observations about how well the scaffolded process operates when input from the scaffolding process is expected but unavailable. This is a rough picture of what such measurements would look like and the empirical work that would have to be performed, but the purpose of this work is merely to present the ideas of ESM in general and leave a robust account of the empirical methodology to further works.

Since we can model process–process scaffolding in this way, we can also apply the analytic tools of graph theory to further characterize the graph and subgraphs, and through such tools we can set down the mind–environment boundary. So far, through process–process scaffolding we have a large network of processes related to each other by scaffolding relationships. This in itself does not delineate which processes are cognitive and which are not. However, this process–process scaffolding model serves as a good foundation for setting a principled mind–environment boundary. My goal in this work is not to formulate the explicit analytics to be used, but to give an account of what such analytics should look like and how they can be used to set down a mind–environment boundary.

Setting down the mind–environment boundary for an individual amounts to identifying some set of processes (a subgraph of the process network) as the mind. Call this the 'mind set.' All processes outside of such a set are, then, the environment. The hard part is identifying such a set. An important advantage of ESM is that it aligns itself with cognitive science, so it is subject to reasonable assumptions that undergird cognitive science. So, we can begin with a reasonable restriction on what the mind set should look like. Namely, the mind set should be continuous: there should not be 'holes' filled with non-cognitive processes in the cognitive network. That is, if p_1 scaffolds p_2 and p_2 scaffolds p_3 where p_1 and p_3 are members of the mind set, then p_2 is also part of the mind set. One of these assumptions on which cognitive science seems to operate is that the mind is continuous. So, to align ESM with cognitive science, I say that the mind set should be continuous.

In principle, we could choose any continuous set to be the mind set for a given minded being, since, as Sterelny says, the demarcation of mind from environment is arbitrary (Sterelny, 2010, p. 480). However, some sets align better with the concept of mind at play in cognitive science. That is, even though in principle we could call any continuous set the mind set, some of these possible sets mesh with the explanatory goals of cognitive science better than others. I claim that to best mesh with cognitive science and thus maximize explanatory power, the mind set should be identified with a highly dense subgraph of the process network, where density is understood as the ratio of cumulative edge weight versus the number of nodes considered.

To see why this maximizes explanatory power, consider what the process network looks like. Presumably, areas around where we would normally approximate minds to be (such as brains) have a high density of scaffolding: there are many relatively highly trusted, integrated, and entrenched processes interacting with each other, and this in general appears to not be true of many other parts of the process network. Instead of thinking that there is anything inherently special about the brain, however, I claim that it is this density of scaffolding itself that is salient—we tend to approximate minds around highly densely scaffolded subgraphs of the process network. This even seems to be part of Sterelny's reasoning for why neuronal memory should be considered part of the the mind while filofaxes should not be. He argued that neuronal memory is much more reliable/trusted than filofaxes and for this reason should be considered to be literally part of the mind, whereas filofaxes are significantly less reliable/trusted and so should not be considered. However, the argument that he makes here and the intuitions behind it are exactly the ones that ESM works on, but in such a way that rigorously sets down the mind–environment boundary. Neuronal memory should be considered to be literally part of the mind because it is part of the highly dense subgraph of the process network, and filofaxes should not be considered to be literally parts of the mind because they are not in a highly scaffolded relationship to processes in the highly dense subgraph of the process network. Thus, in this sense, ESM can sensibly say that neuronal memory is part of the mind while filofaxes are not because the former are highly reliable/trusted (among other dimensions) while the latter are significantly less reliable/trusted. So, ESM accounts for even Sterelny's impression of the salient facts about identifying the mind when his own theory falters. In this way, ESM is able to demarcate the mind in such a way that meshes well with the explanatory goals of cognitive science.

So, density is the salient factor in identifying the mind set, but this does not in itself say how this dense subgraph should be identified. Wherever we delineate the mind set from the environment there are going to be scaffolding relationships crossing this boundary. So, there is no clean division to be made between the highly dense subgraph and the surrounding processes. However, this does not mean that we can't set down such a boundary in a reasonable place. We could use any number of graph theoretical analytics to identify the highly dense subgraph to within some reasonable tolerance. I don't intend to show such analytics here, in part because it would distract from the theoretical points at issue and would require real data to be meaningful, but I take the fact that such mathematically rigorous tools can be applied to identify the mind set to be a major advantage of ESM.

In the place of these rigorous analytics, I offer a rough method for delineating the mind set from the environment that still allows the virtues of ESM to show. The question of whether or not a process is part of the highly dense subgraph becomes less of a question the further one moves to the interior of the dense area. That is, there are processes that are so obviously highly scaffolded that we can take them to be *prima facie* part of the mind set. Portions of the human brain, such as the cortex (which performs processing) and the thalamus (which in many ways functions as a networking hub), are obviously very highly scaffolded, to a degree that few would seriously dispute saying that these processes are part of the highly dense subgraph that is the mind set—these processes are clearly about as scaffolded as the dense subgroup will get. The question of whether or not a process should be considered to be part of the mind set is more salient for processes that are less obviously part of the dense subgraph, but are still more densely scaffolded than the rest of the process network. The rough method I propose is that a process should be included in the mind set if the scaffolding it contributes is not significantly different from the average[2] of the mind set being considered before its inclusion. That is, finding the mind–environment boundary on this rough analytic is a recursive process. Beginning with the obviously dense subset, each process that is being considered for inclusion in the mind set is compared to the set of processes already admitted to the mind set. To preserve the continuity of the mind set, candidate processes for inclusion at any give step of this recursive process are those that are immediately adjacent to the set of processes already admitted to

[2]I am being purposely imprecise about the kind of average needed, since this will be filled in by the chosen rigorous analytic.

the mind set (i.e. they directly scaffold one or more processes already taken to be part of the mind set). If the scaffolding weight this process contributes is significantly less than the degree of scaffolding weight typical in this set of processes already admitted to the mind set, then it should not be included in the mind set. On the other hand, if the weight this process contributes is comparable to the degree of scaffolding typical in this set of processes already admitted to the mind set, then it should be included in the mind set. This is made to exclude processes that make a large step down in terms of scaffolding weight. Processes that yield a small change in scaffolded weight are permissible, but large jumps are prime points to draw the boundary. This rough tool could be supplanted with a rigorous analytic without much difficulty, but this rough tool allows us to see, at least in principle, how ESM works.

For example, consider the weight of scaffolding that my notebook (as opposed to Otto's) contributes to the obviously dense subgraph composed of parts of my brain. The weight of the scaffolding between my notebook and my working memory is significantly weaker than the average scaffolding weight present in this already dense subgraph—for example, compare it to the weight of scaffolding between my neuronal memory and my working memory. Including my notebook in the mind set constitutes adding a significantly weaker scaffold to a set already composed of highly densely scaffolded processes. So, according to ESM my notebook is not part of my mind because there is too significant of a difference in scaffolding weight between what my notebook contributes and the obviously-parts-of-the-mind-set.

What about Otto's notebook? Even in Otto's case, there seems to be a large step down in terms of scaffolding weight between the obviously densely scaffolded subgraph constituted by some of his intracranial processes and his notebook: as Sterelny argues (although not in these terms), it is significantly less reliable/trustworthy than many intracranial densely scaffolded processes. So, we should not consider it to be part of his mind. However, given that Otto has Alzheimer's disease, it may be the case that the average scaffolding weight of his cortex and thalamus, while still greater than the rest of the process network, is lower than people without Alzheimer's disease. This means that Otto's notebook may contribute a less significantly different scaffolding weight than if he did not have Alzheimer's. If this is the case, then according to ESM we should be more apt to say that Otto's notebook is part of his mind since he has Alzheimer's desease than if he did not, since the difference between the scaffolding weight of his notebook and his neural processes is less than if he did not have Alzheimer's disease. And as Sterelny's concerns emphasize, this is the case even if ultimately we still say Otto's notebook is not part of his mind.

The greater the falloff in scaffolding, the easier it is to draw a boundary between the mind set and the environment. That said, ESM has a more difficult time in delineating mind from environment when there is nothing greater than a gradual reduction in scaffolding across a large area of the process network. If there are no large step downs in scaffolding weight, there is no breaking point using the rough method I introduced above. While this is a problem for this rough method, a mathematically rigorous analytic will be able to set down this boundary in just the same way it sets down the boundary for minds with high scaffolding falloff.

At the beginning of this section I stated that ESM and process–process scaffolding would be able to explain everything we can explain in terms of agent–process scaffolding, and now I turn to making good on that claim. We can treat a subgraph as a process unto itself and talk about the processes that scaffold that subgraph. That is, once we've distinguished the mind set from the environment, the agent can just be identified with the mind set and we ask about the processes that scaffold it. While processes may scaffold only a small number of processes within the mind set, we can say that these scaffolding processes scaffold the mind set itself. So, through process–process scaffolding, ESM can talk about agent–process scaffolding in just the same way that Sterelny does by understanding the agent to be the mind set subgraph.

To summarize this section, ESM modifies the scaffolding framework Sterelny sets out. Instead of seeing scaffolding as a relationship between agent and scaffold, ESM sees scaffolding as a relationship between processes. On this view, we can implement graph theory to model all processes on a multidimensional, directed, weighted, transitive graph, called the process network. The feature of this process network that appears associated with minds is scaffolding density. So, setting down the mind–environment boundary amounts to identifying some highly densely scaffolded set of process as the mind and everything else is the environment. Using graph-theoretical analytics along with a few reasonable constraints that ensure ESM identifies the mind as something on par with the goals of cognitive science, we can make this process mathematically rigorous. I take this rigor to be a major advantage of ESM. In the place of such rigorous analytics, which would require much space and real-world data, I offer a rough method for separating the dense subgraph that is the mind set from the rest of the process network. There is some core set of processes that have an obviously highly dense scaffolding weight, to the point that we need not question whether they are in the mind set or not. Other processes should be included in the mind set if the scaffolding weight they would contribute to the mind set does not differ greatly from the average scaffolding density within the mind set before the process under consideration is added. This prevents processes that are significantly less scaffolded than the mind set from being part of the mind. In this way, ESM sets down the mind–environment boundary in a principled and rigorous way that maximizes explanatory power insofar as it meshes well with cognitive science.

3.2 How ESM relates to EM and SM

Now understanding ESM, we should see how it relates to EM and SM. In particular, I show that not only is ESM distinct from both EM and EM, but it has explanatory benefits over each.

3.2.1 ESM vs. EM

Insofar as ESM appeals to scaffolding density, it may appear similar to Clark's coupling criteria and complementariness of EM. However, while they may appear similar, this is not the case. EM used the concept of complementariness to delineate mind from environment. This was to say that so long as a process is integrated with the mind in the right way, it should be considered to be part of the mind. The 'right way' here amounted to the process complementing intracrainal cognitive processes in a highly integrated way. Clark calls this coupling. It may appear that ESM's talk of scaffolding density is just the same as EM's use of the coupling criterion. While the two are on some level similar, there are important explanatory differences. EM does not have a substantial explanation for what it is that makes a process tightly coupled with the mind, whereas ESM does. Moreover, EM does not offer an explanation of how and why such coupling is the salient factor in delineating mind from environment, whereas ESM does. Clark says that processes must be coupled with intracranial processes to count as cognitive, but he does not offer as robust of reasoning as ESM for why coupling is a salient factor in deciding whether or not a process is cognitive. Even further, ESM has all of the explanatory advantages that SM has over EM: EM does not have a good account of how non-cognitive processes can be fundamentally important for cognition, whereas SM and ESM do. For these reasons, ESM is not only distinct from, but is also explanatorially superior to, EM.

3.2.2 ESM vs. SM

Insofar as ESM uses a scaffolding framework and claims all of SM's advantages over EM, it may appear that there is no substantial difference between ESM and SM. However, while they may appear similar, this is not the case. Since ESM uses the process–process scaffolding model, it has a much more articulated understanding of the agent–scaffold relationship than SM. A consequence of this is that ESM has a principled and rigorous way of setting down the mind–environment boundary, whereas Sterelny's SM does not. For these reasons, ESM is not only distinct from, but is also explanatorially superior to, SM.

Moreover, there are interesting questions we can ask using ESM that we can't ask using SM. For instance, with ESM we can address questions about the interconnectivity between cognitive processes: How are they organized? Are there patterns to their structure? Are there separate systems within the mind that only interact at pivotal points, or is there a great degree of connectedness between each process? For humans, these are questions that must be answered by neuroscience. However, while ESM has a framework with which to ask these questions, SM has no resources to address such questions. So, here is another instance where ESM better meshes with cognitive science than SM.

Therefore, not only is ESM distinct from both EM and SM, but ESM has explanatory benefits over EM while also adding explanatory value to SM.

4. CASE STUDIES: APPLYING ESM

Having seen the formal statement of ESM, it will be useful to see it in practice in several different cases. Through doing so, the explanatory power of ESM will become apparent. In this section, I apply ESM to a variety of different kinds of putative minds, both human and radically nonhuman.

4.1 Human Minds

I begin in familiar territory, showing how ESM sets the mind–environment boundary at the skull for modern day humans. I then show how ESM allows this boundary to shift over evolutionary time by exploring how the mind–environment boundary got to be at the skull for modern day humans and the possibility of it expanding beyond the skull in the future.

4.1.1 Modern-day Humans

Above we saw how and why ESM says neuronal memory is part of a modern human's mind while notebooks are not. In doing so, we set down a small part of the mind–environment boundary. This section looks at whether there are strong enough similarities among modern humans to claim a mind–environment boundary in general.

So far, we have looked at ESM in the context of a single mind. Multiple minds exist on the same process network as separate highly densely scaffolded subgraphs with a valley of relatively low-scaffolding processes between them. ESM draws the mind–environment boundary for each of these minds independently, but we can look at the resulting mind sets and see if there are any generalizations we can garner. In doing so, we are looking for general patterns of process inclusion and exclusion from the mind sets of some population, in this case modern humans. To do this, we look for patterns of process inclusion and exclusion across the different mind sets of the population in question. An interesting consequence of this model is that the question ‘What are the boundaries of the modern human mind?’, for example, is an empirical question. The answer to this question comes only after observation of the scaffolding densities of various processes in various people in a quantity great enough to make a population-wide generalization.

With too little data to make a proper population-wide generalization among modern humans, I make such a generalization anyway with the little, anecdotal data available. We are looking for generalizations about what highly densely scaffolded processes modern humans have. First, note what processes most modern humans have in common. While some forms of technology may be relatively common, the only real commonalities between the vast majority of modern humans are intracranial processes. Moreover, while the technologies possessed by modern humans certainly scaffold cognition, the scaffolding weight they contribute is far less than the average density of intracranial cognitive processes. So, extracranial processes fail twofold to be cognitive for modern humans generally: they are neither common nor highly scaffolded relative to the obviously highly densely scaffolded subgraph. If a single modern day human were to have some process that contributes scaffolding weights comparable to intracranial processes, then such a process would be part of his mind, but this isolated case would not impact the generalization that, for modern humans in general, cognition is limited to neuronal processes.

For further examples, let's return to the neuroprosthetics that motivated this work. Cochlear implants, deep brain stimulation, as well as the devices created by Berger et al. and Andersen et al., appear to contribute scaffolding with weights not significantly different neuronal processes, and thus should be considered to be parts of the minds of the people who have them, according to ESM. Again, the number of modern humans that have these devices is small, so such devices do not change the claim that in general, modern human cognition is limited to neuronal processes.

4.1.2 Past Minds

Presumably, human cognition did not begin in its current form—the modern human's cognition is built upon successive expansions and developments in cognition. That is, over evolutionary time the hominid mind has expanded, and we should have an account of how this happened.

Sterelny is certainly interested in explaining this fact, and goes on to give an explanation about how hominid memory may have evolved from simple detection systems. He begins with simple detection systems, such as those present in bacteria and plants. These systems allow an organism to monitor their environments and respond in very predictable ways (Sterelny, 2003, p. 14). Through these systems, an organism monitors for a single environmental cue, then always responds in the same predictable way. Because of this predictability, these systems are easily exploited. For example, ants have simple detection systems that monitor for parasites and colony invaders. Ants respond to specific chemical compounds to differentiate other ants of the colony from invaders: so long as these chemical cues are present, the ants treat an organism as part of the colony. This simple system is exploited by a number of species of beetles that need only mimic these specific chemical cues to gain free roam over the colony and exploit the colony's resources (Sterelny, 2003, p. 15).

One evolutionary response to this exploitation of simple, single-cue detection systems is to incorporate more cues and develop what Sterelny calls robust tracking systems. That is, to keep parasites or predators from exploiting the tightly-bound stimulus–response behavior of a single cue, an organism uses multiple cues to determine an action. Since more cues have to be met in order for a predator to elicit a predictable response, the predator must put forth more effort to take advantage of predictable behaviors. Also, having redundant cue systems allows organisms to compensate in variable environments. For example, bees have at least three cue systems by which they can navigate from hive to food source: landmarking, sun direction, and sunlight polarization. By having these multiple cue systems, bees are able to compensate for one system being unavailable due to variable environmental conditions, such as a cloudy day or removal of landmarks (Sterelny, 2003, p. 24).

The important difference between simple detection systems and robust tracking systems is response breadth. Simple detection systems have a narrow response breadth, insofar as a single cue is tightly coupled with a predictable behavior. Robust tracking systems have a broad response breadth, since a single cue is not tightly coupled with any specific behavior. According to Sterelny, the upper limit point of response breadth are decoupled representations: robust tracking systems that have a very broad response breath to the point that connections between cue and response are very loose.

According to Sterelny, decoupled representation in hominids may have evolved from the decoupling of spatial representations (Sterelny, 2003, p. 41). Many animals have good spatial memory: the storage of information about their immediate environment. Some animals store this information as procedural maps, such as representing the path from A to C with information about how to go from A to B and from B to C. In this case, if B is removed or the organism is unable to detect it, the animal will not be able to travel from A to C. On the other hand, other animals have a more decoupled representation of the information, whereby they store a cognitive map of the environment between A and C and are able to choose any path between the two (Sterelny, 2003, p. 41-42). The latter has the adaptive advantages of decoupled representation, and may have been one step on the way to hominid decoupled representation.

The next step, for Sterelny, is from decoupled spatial maps to decoupled social maps, where these social maps are "a plausible origin of belief-like representations" (Sterelny, 2003, p. 51-3). For hominids, social interactions exert evolutionary pressure, and this requires robust social recognitional capacities and social memory (Sterelny, 2003, p. 53). Either on the same systems as decoupled spatial representations or analogous ones, hominids developed decoupled social representations with which to navigate the complex social environment. According to Sterelny, these decoupled social representations are also a plausible origin of belief-like representations. This is expressed in his version of the Social Intelligence Hypothesis, which claims that folk psychology and the concept of beliefs evolved as the adaptation to operating in a society of organisms with behaviorally decoupled representations (Sterelny, 2003, p. 56-7). That is, given the demands of society, one must also behave in certain ways to avoid bad outcomes. Knowing how to behave well in society requires anticipating the behaviors of others. This is difficult because hominid society is composed of individuals that already have decoupled representation in one way or another. This makes individuals' behaviors unpredictable, since behaviors are not tightly coupled with given cues. An adaptation to overcome this difficulty is to develop a folk psychology that uses decoupled, belief–like representations to perform the function of mind reading. Attributing beliefs to individuals of society makes behavior prediction more accurate, and more robust in cases where specific cues are missing. Thus, Sterelny has an account of the evolution of modern human cognition from simple detection systems.

It is not clear that SM has a good way of addressing this changing boundary. At best, we could say that since Sterelny is vague and noncommittal about the details of the boundary (given his explanatory goals), his theory is not incompatible with a changing boundary. While SM may be just vague enough to work, ESM can do better. More than being merely consistent with a changing boundary, ESM has a robust account of how and when the boundary of hominid cognition changes over evolutionary time. When environmental pressures selected for increasing decoupling of representations in systems like memory, this selection increased scaffolding weight. That is, as something like a robust tracking system evolves, the processes that undergird it gain scaffolding density. These processes may eventually rise in scaffolding density to the point that they are comparable with already highly scaffolded processes, and thus become part of the organism's mind. So, over evolutionary time cognitive adaptations go from having low scaffolding weight to being highly densely scaffolded. ESM gives analogous accounts of the development of space and social maps, and eventually gives the same account of the evolution of folk psychology. In this way, ESM gives a detailed account of how cognitive boundaries expand with the widespread integration of robust tracking systems. Other cognitive advancements can be accounted for in an analogous way. So, ESM gives a robust account of how cognitive boundaries have expanded over time.

4.1.3 Future Minds

Continuing the account ESM gives about how human cognition evolved in the past to its present day boundaries, ESM can also give a robust account of how human cognition may continue to expand in the future. As we have seen, some modern humans have cognitive boundaries extending beyond neurons, according to ESM. To be able to say that for humans in general the boundary of cognition extends into such devices, according to ESM it would require only that a sufficient proportion of the human population had such devices in highly densely scaffolded ways to justify the generalization. So, whereas SM does not offer a robust account of what it would take for cognition to extend beyond the skull, ESM allows for cognition to extend beyond the skull in the future and has a principled way of determining if and when this happens.

4.1.4 Synchronic Group Diversity

So far I have discussed ESM at the level of the individual and the species level. However, ESM can equally well apply to any grouping in between these levels. A consequence of this is that ESM can account for synchronic group diversity in cognitive boundaries. According to evolutionary theory, different groups of people separated by barriers, whether they be geographical, cultural, or status-based, can have divergent evolutions given enough time. One of these factors that may change between different groups of people is the cognitive boundary. ESM allows us to easily account for cognitive boundary differences between individuals posessing the same kind of mind by making two generalizations. While we may be able to generalize about hominids as a whole only to a limited extent, different subgroups can have more extensive boundaries. For example, while cognition is limited at the skull for modern humans, the boundary of cognition extends beyond this for the group of humans that have cochlear implants. So, ESM can not only make claims about individuals and hominids as a whole, but can also make claims about subsets of humans, all while using the same rigorous process.

4.2 Non-Human Minds

Since one of the virtues of ESM is that it meshes well with cognitive science, and since cognitive science studies non-human entities in addition to human cognition, it would do ESM well to be able to account for non-human cognition. ESM can indeed account for non-human cognition: if an entity is accepted as being minded, whether human or non-human, ESM sets down the boundary of that entity's mind in the same rigorous way. For other animals, this is fairly straightforward and I forgo a discussion of these cases to pursue an account of more challenging kinds of minds. I show how ESM can set a mind–environment boundary for radically non-human minded beings using IBM's Watson supercomputer as an example.

Watson, the computer to win *Jeopardy!* in 2011, is the product of IBM's DeepQA Research Team, which aims to create the first open-domain natural language question-answering system (Ferrucci, 2012, p. 1). IBM brands Watson as a "cognitive system" (*IBM Research: Watson*, n.d.). Suppose we take this literally and grant that Watson is, in some relevant way, minded—that is, suppose we accept some philosophy of mind that takes Watson to be a cognitive, minded being.[1]

[1] If one finds such a supposition to be impossible, then this example application of ESM holds no weight. However, this does not tell against ESM, since one's issue is with the philosophy of mind rather than ESM, since ESM is not a tool for identifying minded beings but only setting cognitive boundaries for organisms already supposed to be minded.

We can then apply ESM to Watson and delineate Watson's mind from its environment, but we must first look at the process network as it pertains to Watson. That is, we must first study Watson's anatomy and delineate a highly densely scaffolded subset of the process network. Watson is designed to be able to dynamically interpret unstructured data, including natural human language. It relies on the DeepQA architecture, which in many ways mirrors the functional work flow of humans playing *Jeopardy!*. DeepQA begins by parsing, interpreting, and understanding the input question. It begins by guessing what sort of answer the question requires and parses the question grammatically (Lally et al., 2012; McCord et al., 2012). DeepQA then accesses its stored memory to search for candidate answers. Importantly, Watson is self-contained; it does not merely perform a Google search for information. Like all *Jeopardy!* contestants, it has to learn, parse, and store information before the contest. The research team loaded Watson with obvious sources of information, like encyclopedias, reference books, and Wikipedia and a method of incorporating other web resources (Chu-Carroll, Fan, Schlaefer, & Zardrozny, 2012). This unstructured information is then grammar-parsed and processed into structured information, including a source-reliability index (Fan et al., 2012). Some of this processed information takes the form of commonsense axioms, such as "books are found on shelves," "people visit museums," and "candidates win elections" (Ferrucci, 2012). In this way, Watson accumulates a corpus of self-contained knowledge, similar to the way in which humans learn information from a source, process it, and store that processed information in on-board hardware (i.e. memory neurons) for quick access later. DeepQA

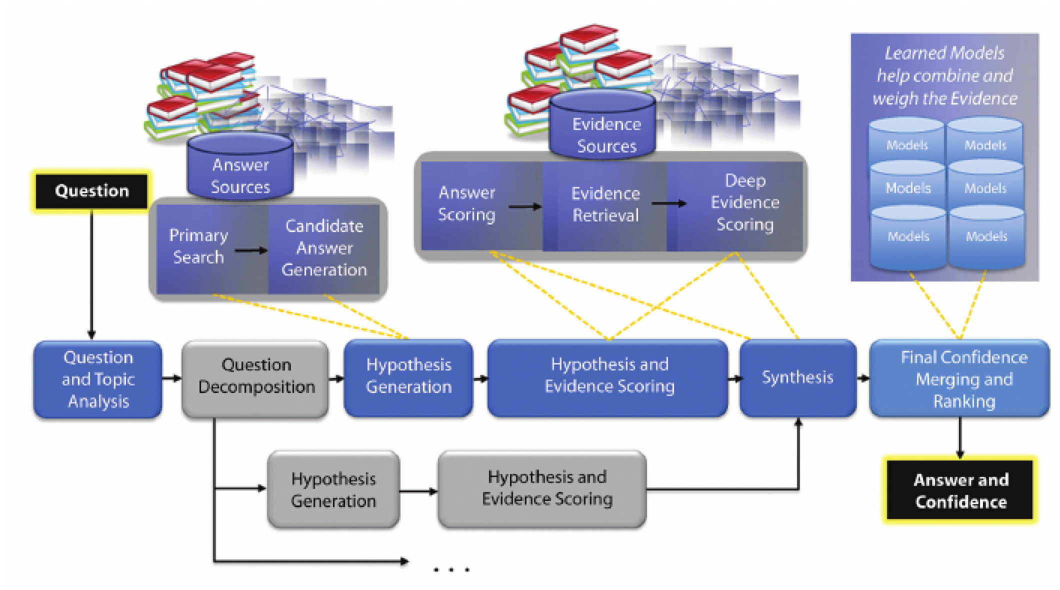

Fig. 4.1. DeepQA architecture. Reproduced from Ferrucci (2012, p. 5).

uses this stored information to generate candidate answers to the question, ranked by confidence level (Chu-Carroll, Fan, Boguraev, et al., 2012). It then searches its corpus of axioms, looking for evidence for or against each candidate and updates its confidence levels (Murdock et al., 2012). This DeepQA architecture is then wrapped in the Watson architecture, which adds *Jeopardy!* strategy, such as what confidence threshold should be met before ringing in and how much to bet on *Double Jeopardy* (Ferrucci, 2012). See fig. 4.1 for a full schematic of the DeepQA architecture. The DeepQA Research Team sees applications for DeepQA beyond game shows, including suggesting medical diagnoses (Ferrucci, 2012).

Since Watson is, putatively, minded, we can ask where the boundary between Watson's mind and its environment is. For instance, is the internet part of Watson's mind? ESM gives us a way of giving such a boundary. It says that we should expect there to be some highly densely scaffolded set of processes around where we would expect Watson's mind to be, and then we can determine whether or not a given process is cognitive based on whether the scaffolding weight it offers is comparable to the scaffolding density already within this obviously dense set.

With the anatomy of Watson's architecture in mind, we can see that there is indeed such an obviously densely scaffolded set, which includes its 90 networked IBM Power 750 servers containing 2,800 POWER7 cores, 16 Terabytes of RAM, and 4 Terabytes of disk space (Deedrick, 2011). These processes interact in highly trusted, individualized, and entrenched ways. They are trusted because each process relies on its input from other sectors without verifying the accuracy of the data. One point where this may seem to not be true is in the "Hypothesis and Evidence Scoring" part of the DeepQA architecture. This step takes proposed hypotheses from the "Hypothesis Generation" module and checks them against a database of information to assign probabilities of each hypothesis being true. This may initially appear like mistrust of the hypothesis generator, since the scoring module does not accept the hypotheses as already true, but this is a misunderstanding of the kind of trust relevant for scaffolding. The kind of trust under consideration here focuses on two components: the process giving the input is 1) functioning how it should, and 2) is readily available. We can see that both of these conditions are met: the scoring module trusts that the hypothesis generator is functioning as it should to produce plausible hypotheses, and the generator is always available whenever the scoring module needs its input. So, even in this case, these processes of Watson are highly trusted. This set of processes also has a high density of individualization. Watson's processors, RAM, and disk space are used solely by Watson. These processes also have a high density of entrenchment. Each processor and piece of RAM and disk space is filled with information and programming sequences, so that simply replacing one of these components with a new one will result in system errors.

On the other hand, resources like its power supply, cooling mechanisms, and even the internet are significantly less highly scaffolded. While necessary, individual power supply lines and cooling mechanisms can be replaced without affecting Watson's functioning. More interesting, perhaps, is the internet. While Watson does depend on the internet as a source of information to be stored in its database, the internet is not highly scaffolded for Watson. While the information that goes into its database has been preprocessed manually, Watson does not trust any single piece of information gathered from the internet. When using this information, Watson always verifies information it finds against other sources and preset axioms. So, the internet does not rank very highly for Watson on the trust scaffolding dimension. Obviously, the internet is not highly individualized for Watson—Watson is far from the only organism to access and use information on the internet. Moreover, the internet is not highly entrenched for Watson. While Watson requires information from the internet at first, during its actual operation Watson is completely disconnected from the internet. So, the internet is not highly scaffolded for Watson. Thus, ESM claims, we should set the boundary of Watson's mind at its processors, RAM, and disk space and say that power cables, cooling mechanisms, and the internet are all parts of its environment.

5. CONCLUSION

Neuroprostheses make the question of where one's mind ends and the environment begins a pressing question for modern humans. We are at the precipice of living in a time when these devices become widespread, and whether or not these devices are truly cognitive will be an important philosophical, scientific, and political issue. Ultimately, this work proposes ESM as an answer to this question.

ESM is a development of two previous answers to the question of how to set down the mind–environment boundary, namely EM and SM. EM (defended by Clark and others) proposed to treat any process that sufficiently complements intracranial cognitive processes as cognitive processes themselves. SM (defended by Sterelny) uses scaffolding to explain how extracranial process can be important for cognition yet not cognitive processes themselves. Clark and Sterelny argue over which theory has the superior explanatory power: Clark argues that EM is the only theory that can adequately explain the importance of extracranial processes to cognition. Sterelny argues that not only does SM account for the importance of extracranial processes, but it can explain the entire spectrum of reliance on non-cognitive tools through scaffolding.

I agree with Sterelny that SM has more explanatory power than EM. However, while SM fits Sterelny's explanatory goals, it does not address the question of where the mind–environment boundary actually lies.

My proposal, ESM, uses a modified version of Sterelny's scaffolding to set down the mind–environment boundary. Instead of looking at scaffolding as a relationship between agents and scaffolds, ESM looks at scaffolding as a relationship holding between processes. On this process–process conception of scaffolding, we can model processes using graph theory, where we let nodes represent processes and the connections between them represent the scaffolding relation. In keeping with the concept of scaffolding, this resulting graph is directed, weighted, and transitive, with transitive weights taking the minimum value of the steps in the transitive chain. Ultimately, this model allows us to set down the mind–environment boundary in a principled and rigorous way that also meshes with the explanatory aims of cognitive science in general.

I have argued that the key factor in identifying the mind is scaffolding density. That is, we should identify the mind with the set of highly densely scaffolded processes. Since the network of processes is continuous, there is no clean break to be made between mind and environment. However, since ESM models scaffolding using graph theory, we have at our disposal the rigorous analytics of graph theory to make such a determination. How exactly such determinations are made may differ according to one's explanatory aims, but I offered a rough method by which to determine whether a process is densely scaffolded enough to be included in the mind of an organism. We can generally identify some core set of processes as obviously highly densely scaffolded, and so obviously part of the mind. The question of whether or not a process is part of the mind is more salient for processes outside of this set. For such processes, this rough method says that we should include them in the mind of an organism if the scaffolding that they contribute does not differ significantly from the average density of the obviously dense core.

I then showed how ESM applies to several different classes of minded beings, including past, present, and future humans as well as radically non-human minds such as IBM's Watson. Ultimately, ESM sets the mind–environment boundary for modern humans at roughly the same place Sterelny does: the skull. However, ESM adds to SM the ability to give a robust and principled account of where the boundary of cognition should be set down. As a result, it can account for shifting cognitive boundaries as well as non-human boundaries and varying boundaries between different members of the same group. In addition, it meshes better with cognitive science both in terms of explanatory scope and ability to integrate mathematical modeling based on quantitative data.

Therefore, ESM presents a rigorous and principled way of setting down the mind–environment boundary in such a way that maximizes explanatory power over both EM and SM.

REFERENCES

REFERENCES

Andersen, R. A., Burdick, J., Mulallam, S., Persaran, B., & Cham, J. (2004). Cognitive Neural Prosthetics. *Trends in Cognitive Sciences*, *8*(11), 486–493.

Andersen, R. A., Hwang, E. J., & Mulliken, G. H. (2010). Cognitive Neural Prosthetics. *Annual Review of Psychology*, *61*, 169–90.

Berger, T. W., Song, D., Chan, R. H. M., Marmarelis, V. Z., LaCoss, J., Wills, J., ... Granacki, J. J. (2012). A hippocampal Cognitive Prosthesis: Multi-Input, Multi-Output Nonlinear Modeling and VLSI Implementation. *Transactions on Neural Systems & Rehabilitation Engineering*, *20*(2), 198–211.

Chu-Carroll, J., Fan, J., Boguraev, B. K., Carmel, D., Sheinwald, D., & Welty, C. (2012). Finding needles in the haystack: Search and candidate generation. *IBM Journal of Research and Development*, *56*(3.4).

Chu-Carroll, J., Fan, J., Schlaefer, N., & Zardrozny, W. (2012). Textual resource acquisition and engineering. *IBM Journal of Research and Development*, *56*(3.4).

Clark, A. (2008a). Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind? *Philosophy and Phenomenological Research*, *76*(1), 37–59.

Clark, A. (2008b). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* Oxford UP.

Clark, A. (2010). Memento's Revenge: The Extended Mind, Extended. In R. Menary (Ed.), *The Extended Mind.* MIT Press.

Clark, A., & Chalmers, D. J. (1998). The Extended Mind. *Analysis*, *58*(1), 7–19.

Deedrick, T. (2011). It's Technical, Dear Watson. *IBM Systems Magazine*. (http://www.ibmsystemsmag.com)

Fan, J., Kalyanpur, A., Gondek, D. C., & Ferrucci, D. A. (2012). Automatic knowledge extraction from documents. *IBM Journal of Research and Development*, *56*(3.4).

Ferrucci, D. A. (2012). Introduction to "This is Watson". *IBM Journal of Research and Development*, *56*(3.4).

Hampson, R. E., Song, D., Chan, R. H. M., Sweatt, A. J., Riley, M. R., Gerhardt, G. A., … Deadwyler, S. A. (2010). A Nonlinear Model for Hippocampal Cognitive Prosthesis: Memory Facilitation by Hippocampal Ensemble Stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(2), 184–197.

Hampson, R. E., Song, D., Opris, I., Santos, L. M., Shin, D. C., Gerhardt, G. A., … Deadwyler, S. A. (2013). Facilitation of memory encoding in primate hippocampus by a neuroprosthesis that promotes task-specific neural firing. *Journal of Neural Engineering*, *10*, 1–16.

IBM Research: Watson. (n.d.). (http://ibm.co/1emj6Xn)

Kirchhoff, M. D. (2012). Extended cognition and fixed properties: steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences*, *11*, 287–308.

Lally, A., Prager, J. M., McCord, M. C., Boguraev, B. K., Patwardhan, S., Fan, J., ... Chu-Carroll, J. (2012). Question analysis: How Watson Reads a Clue. *IBM Journal of Research and Development*, *56*(3.4).

McCord, M. C., Murdock, J. W., & Boguraev, B. K. (2012). Deep parsing in Watson. *IBM Journal of Research and Development*, *56*(3.4).

Murdock, J. W., Kalyanpur, A., Welty, C., Fan, J., Ferrucci, D. A., Gondek, D. C., ... Kanayama, H. (2012). Typing candidate answers using type coercion. *IBM Journal of Research and Development*, *56*(3.4).

Sterelny, K. (2003). *Thought in a Hostile World.* Blackwell.

Sterelny, K. (2004). Externalism, Epistemic Artefacts and the Extended Mind. In R. Schantz (Ed.), *The Externalist Challenge.* Walter de Gruyter.

Sterelny, K. (2007). Cognitive Load and Human Decision, or, Three Ways of Rolling the Rock Uphill. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind: Volume 2: Culture and Cognition.* Oxford UP.

Sterelny, K. (2010). Minds: extended or scaffolded. *Phenomenology and the Cognitive Sciences*, *9*, 465–81.

Sutton, J. (2010). Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process. In R. Menary (Ed.), *The Extended Mind.* MIT Press.

Sutton, J., Harris, C. B., Keil, P. G., & Barnier, A. J. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, *9*, 521–60.

Zeng, F.-G. (2004). Trends in Cochlear Implants. *Trends in Amplification*, *8*(1).

Zibetti, M., Merola, A., Rizzi, L., Ricchi, V., Angrisano, S., Azzaro, C., … Lopiano, L. (2011). Beyond Nine Years of Continuous Subthalamic Nucleus Deep Brain Stimulation in Parkinson's Disease. *Movement Disorders*, *26*(13), 2327–2334.