

Evolution & Experience

By Farhan Lakhany*

One way of making sense of experiential states is via an evolutionary explanation. To do this, we need to clearly explicate the adaptive role which experiential states play in the organisms which have it. In his works, Peter Godfrey-Smith provides two different ways of understanding experiential states as evolutionary adaptations and calls the two views the “transformation view” and the “latecomer view”. Godfrey-Smith adopts the transformation view, and, in this article, I defend the latecomer view. This paper is broken up into six sections. In the first section, I provide an overview of the transformation view and the latecomer view. In the second section, I provide a more detailed analysis of the latecomer view, highlight its connection to the global workspace theory of consciousness and provide an argument for it. In the third section, I provide a response on behalf of Godfrey-Smith to my argument and provide two arguments for why we should prefer the transformation view to the latecomer view. In the fourth section, I argue against the transformation view by claiming that, unlike the latecomer view, it does not provide an adaptive role for experiential states. In the fifth section, I attempt to provide, on behalf of Godfrey-Smith and the transformation view, an adaptive role for the transformation view but argue that the response fails. In the sixth section, I respond to the two arguments articulated in the third section and show how they can be overcome. I conclude by discussing the upshot of the arguments presented in the article.

Keywords: *Evolution, experience, global neuronal workspace, primordial emotions, latecomer view, transformation view*

Introduction

Unless we claim that all biological organisms¹ have experiential states² or that no biological organisms have experiential states, we seem consigned to the messy middle ground in which some biological organisms have experiential states, and some do not. Human brains (comprised of roughly 86 billion neurons) would certainly fall in the former category while *c. elegans* worms³ (comprised of roughly 300

*Professor, University of Nebraska Omaha, USA.

¹For this paper, I will define “biological organism” as, “An individual living thing that can react to stimuli, reproduce, grow, and maintain homeostasis” (Biology-online 2014). This definition is quite broad and arguably would include in its purview entities like bacteria and viruses.

²For this paper, I will be understanding an experiential state as a state of an information processing system with experiential properties. Further, I am understanding the presence of experiential states in a binary fashion; a state is an experiential state or it is not. Further, I use the terms “experiential”, “subjectivity” and “subjective experience” interchangeably. There are semantic differences between these terms but they will not be relevant for the forthcoming account.

³I do not take this claim to be uncontroversial as drawing the line for what appear to be obvious cases for what should and should not be considered conscious is a hotly debated topic. For example, there

neurons) would arguably fall in the latter (Herculano-Houzel 2009; White et. al., 1986). If the previous sentence is correct, any attempt to explain the existence of experiential states in biological entities must address why it is that certain biological entities have these states and why others do not. One way in which we might explain the existence of experiential states is to provide an evolutionary account⁴. In this paper, I will attempt to support one such account and, in doing so, will take it as granted that experiential states, similar to wings and opposable thumbs, are properly understood as an evolutionary adaptation. Peter Godfrey-Smith provides two ways of understanding experiential states as evolutionary adaptations by distinguishing between the “transformation view” and the “latecomer view” with his preferred formulation being the former. For this paper, I defend the latter view from perceived deficiencies by closely analyzing Peter-Godfrey Smith’s critique of it and his own account of the emergence⁵ of experiential states. The paper will proceed in the following way: first, I articulate three assumptions that I’m taking on for this paper and note the two positions on the evolution of experiential states as highlighted by Godfrey-Smith; second I provide a more detailed overview of the “latecomer view” by highlighting the motivations for it and provide one argument in its favor; third, I provide a response on behalf of Godfrey-Smith to my argument for the “latecomer view” and proceed to explicate and provide two arguments for the “transformation view”; fourth, I argue against the “transformation view” by showing how it seems to not provide a functional role for experiential states in the way that the “latecomer view” does; fifth, I attempt to fill this gap by providing, on behalf of the “transformation view” an argument and proceed to argue that this response fails; and finally, sixth, I respond to both of the earlier arguments that Godfrey-Smith leveled against the

is interesting work done by Paco Calvo to argue for the possibility of plants (despite their lack of a central nervous system) having experiential states on the grounds that they seem to exhibit the necessary behavior that is sufficient for the attribution of consciousness to other organisms (Calvo 2017). With that said, I do follow the lead set out in Calvo and assume that one of the necessary conditions for the attribution of experiential states is *anticipation* or *goal-directed* behavior and it appears that worms seem to lack such capacities (Calvo 2017).

⁴There are some which will find such an account to be dubious because of broad worries that any evolutionary account will, in virtue of its more than likely being a physicalist account, run into the explanatory gap when it comes to experience (cf. Levine 1983). I have two responses to this worry: first, one might avoid such a worry by understanding experiential states in physicalist terms (for example, one might attempt to identify it with certain syntactic properties of a (or set of representational) representational state of an information processing system); second, an evolutionary account does not *necessitate* one to accept a reductive physicalist account. To see this, we need only remember that all evolutionary accounts provide is an account of how a certain *function* (or proto-function (the terminology here will depend on metaphysically loaded we understand the notion of “function”)) was “picked” by a process of natural selection. It *might* be that what was “selected” was a non-physical property which supervenes on the physical state of an organism. There are, of course, a whole different set of problems which arise when taking the supervenience route, but this is a live option and much of what is said here could be paired with a non-reductive physicalist account. That said, I am implicitly assuming that experiential states *cause* an organism to act in some way and leave it to the reader to cash out how this mental causation might work.

⁵Emergence is not intended to be understood in a metaphysically loaded sense here; by emergence I mean to be highlighting an emergent *function* in the same way that the ability to fly might be understood as an emergent function of having wings (though again, if one wants to make the function dependent on some non-physical property, this is open for the individual to do so).

“latecomer view” and show how they can be overcome. With all that said, I proceed to the first part.

I

Confronted with the fact that some biological organisms have experiential states, and some do not, an ideal explanation would provide necessary and sufficient conditions and explain why those conditions need to hold. That project, while laudable, is in a way a step ahead of the main interest of this paper. Here, I take on a more basic question that helps to lay the conceptual groundwork for locating the strict definitional criteria: I explore where, in the evolutionary history⁶ of biological organisms, we might find the origins of subjective experience by attempting to understand the function of experiential states.

To start tracking down the evolution of experiential states, we might begin by attempting to articulate what these experiential states are and what they do. This is clearly a massive undertaking well beyond the scope of this paper so I will make some assumptions to wrestle these questions into a manageable state:

- 1) I make an assumption about what experiential states are: I will consider the experiential states of an organism as functional states⁷ (i.e., they affect the behavior of the organism) and the way they end up affecting the functioning of the organism is by modulating the way in which organism processes information⁸.
 - 1a. I assume that experiential states are either supervenient on or identical to a global state of information processing.⁹
- 2) Second, I assume the causal closure of the physical.
- 3) Third, I make an assumption about why we have these states: I assume that having experiential states provides an adaptive advantage and that

⁶It is important to note that while the commonplace claims of being “more” or “less” evolved might encourage a linear understanding of the evolutionary history of biological organisms, this assumption is, strictly speaking, incorrect. Evolutionary history is better understood as a tree where all biological organisms as sharing a common “trunk” from which the different branches evolved down their own paths. When I speak in this paper of “later” or “earlier”, this is to be understood relative to different branches with “later” being understood as more functionally complex and “earlier” being understood as less functionally complex.

⁷Admittedly, this does little to whittle down the ontological field; I do not take a definitive stand on this point on purpose and all I mean to be doing is ruling out epiphenomenal accounts of experiential properties.

⁸I do not attempt to define “information” in this paper though I do sympathize with the frustration a reader might feel in relation to this philosophically controversial notion.

⁹The necessity of such a delineation between information states might be challenged; one might plausibly ask why we cannot simply identify information states with experiential states. In response one need only note that information processing in biological organisms seems to predate the existence of experiential properties. To see this, we can consider the fact that *C. elegans* worms are able to process information about its’ oxygen levels and yet, as I will argue, they are a case of a biological organism that does *not* have experiential properties (Oda et. al., 2017). If it is correct that they do not have such states (and I will argue they do not as the paper continues), then we need to provide an account for why *some* information states have experiential properties and others do not.

advantage is a byproduct of the increased repertoire of behavior that the organism can either actively engage in or plan future behavior by simulating those possible behaviors internally¹⁰.

With that stated, I will begin by orienting my enquiry to align with Peter-Godfrey Smith in his book *Other Minds: The Octopus, The Sea, and The Deep Origins of Consciousness* by asking whether experiential states are recent evolutionary adaptations or older evolutionary adaptations. In *Other Minds*, Godfrey-Smith refers to the view that experiential states are a relatively early evolutionary adaptation as the “transformation view”¹¹ and the view that experiential states are relatively late evolutionary adaptation as the “latecomer view”¹² (Godfrey-Smith, 2018).

Of these two views, Godfrey-Smith believes the TFV to be correct. I disagree and instead adopt the LCV for reasons that I will address in this paper. But before I get to this, some more stage setting is required in which I specify both the TFV and the LCV in more detail. In doing this, I will also provide arguments for why an individual might adopt each view. I start with the LCV.

II

To investigate experiential states, we can begin by studying a rather unique (and self-aggrandizing) member of the great ape family that we can be sure have experiential states: humans. We know that *we* have experiential states and, coupled with the assumption that they modulate the way in which we process information, we can attempt to understand how these experiential states allow us to act in ways that we otherwise could not. As neuroimaging technology has improved, this project has turned out fascinating results with one theory in particular, the global workspace theory of consciousness, steadily gaining adherents. One relevant aspect of this theory is its reliance on certain complex cognitive mechanisms and styles of information processing. Due to the complexity of these mechanisms, the organisms which have them appear late in the evolutionary history of organic life¹³ and as a result, individuals which hold the global workspace theory of consciousness naturally fall into the LCV of experiential states (Godfrey-Smith, 2018). Thus, to understand the appeal of the LCV (and later, the argument I give in its favor), it is worth knowing a bit about the ‘global workspace’ view and its commitments.

Understanding consciousness as a ‘global workspace’¹⁴ was pioneered by work from Bernard Baars in 1980s but has been more recently spearheaded by Stanislas

¹⁰“Internally” is not to be understood in any metaphysically loaded sense; all I mean is that the organism can process the different behaviors and make informed guesses about the possible ramifications of those behaviors.

¹¹From here on referred to as the “TFV”.

¹²From here on referred to as the “LCV”.

¹³Exactly which beings fall on either side of the divide is not discussed here; a rough sketch will include human beings and likely many different species of mammals and birds (Godfrey-Smith, 2018).

¹⁴Elaborated as a “global neuronal workspace” by Dehaene et al. (Dehaene, 2014). For this paper, I will refer to the “global neuronal workspace” as the “GNW”.

Dehaene in collaboration with Lionel Naccache and Jean-Pierre Changeux (2014)¹⁵. To grasp the GNW view of consciousness, we need a very simplified, high-level understanding of brain function. One plausible way of understanding the early function of brains¹⁶ is as a basic information processor which took in information from environmental stimuli as an input and produced some behavior as an output (Braitenberg, 1987). From here, brains have evolved into ever more sophisticated information processors and handlers¹⁷. This processing/handling occurs in very specialized¹⁸ circuits¹⁹ that allow for quick, dependable, responses to environmental stimuli²⁰.

These circuits are often individuated along functional lines²¹ that lead to neurally localized²² activation (e.g. the Broca's area in the production of speech). This dedicated circuitry has the upside of being extremely fast while sacrificing flexibility in the type of information that can be processed in these regions. These (relatively) functionally isolated islands of activity works well for organisms with a (relatively) simple repertoire of stimulus-response behaviors and an environment that offers few survival challenges. In these scenarios, flexible processing would not necessarily provide an adaptive edge in a race for survival.

However, once we increase the complexity of the organism by adding more local functions into the biological system and complicate the environment, both a new possibility and new challenge arises. The new possibility is a combinatory possibility: with so many possible actions arising as a byproduct of localized information processing, more complex behavior becomes possible as different combinations of local functions might be strung together to create a more sophisticated, global action. This capacity would allow the organism to respond more effectively to environments characterized by a low degree of predictability as the organism can tap into a vastly expanded set of possible behaviors in response to unforeseen contingencies.

But this synchronized activity brings its own host of problems: namely, developing systems to actually synchronize the activities of the subsystems. If this hurdle can

¹⁵For the sake of specificity, I will be relying on Dehaene's exposition of the global neuronal workspace from his book, *Consciousness and the Brain*.

¹⁶Strictly speaking, these might better be classified as "nervous systems" and not full blown "brains". As it is not relevant to this paper, I am slightly loose with the terminology but it is worth noting that the terms are not necessarily coextensive.

¹⁷By "handling", what I mean to pick out is the capacity to not only process information but to collate and route information to different processors in the brain.

¹⁸"Specialized" is to be understood relative to the make up of the rest of the brain. Indeed, some recent research has shown the importance of multiple brain networks working in conjunction to achieve certain functions (Jasanoff, 2018).

¹⁹For this paper I simply use the notion of "circuit" but it is worth knowing that this runs roughshod over some distinctions; circuits can be smaller or larger and, when large enough, can be considered a "micro-network" inside of a larger brain network (Pulvermüller, 2014).

²⁰Of course, as alluded to previously, this circuitry can also be used purely internally and in the absence of external stimuli by simulating different scenarios.

²¹Providing the necessary and sufficient conditions for certain cognitive functions is disputed and, for this paper, I will be relying on a relatively intuitive notion of "function". Further, I will not be specifying the different functions as it is not directly relevant to the main thesis of this paper.

²²"Localized" is to be understood as relatively localized; neuronal activation is not necessarily as clean cut as it might ideally be "spillover" neuronal activation does seem to occur (Huang, 1998).

be overcome, the beginnings of a truly game-changing ability can be found: the ability to exercise top-down control to marshal the resources of the organism to act in the service of a goal that the local functions on their own could not achieve. In this ability, we see the seeds of the GNW.

At its heart, the GNW is a much more sophisticated functional solution to achieving goal-directed, synchronized, complex action. Very roughly put, the GNW can be understood as an information buffer²³ which is sourced from the various information processing centers of the brain which is then made available for high-level decision making²⁴. This last aspect is crucial; recall from before that one of the solutions to dynamic, adaptable behavior is the exercise of top-down control to marshal the resources of an organism to act in service of a goal that would otherwise not be achievable locally.

The non-local functions that the GNW allows for are certain, higher-order abilities: collating, organizing and integrating multiple sources of information processed by a system *for another part of the system*. Conscious states, on this account, serve as a kind presidential briefing for another information processing aspect of the organism²⁵.

This is all interesting from a neuroscientific point of view; what makes the GNW *philosophically* interesting is that its presence appears to directly correlate with the presence of conscious states in human beings. This finding has led to very exciting advancements in theorizing about both the function and explanatory role of experiential states that can be rigorously tested in an experimental setting.

By itself, the information presented above points in the direction of thinking that the GNW might best describe human consciousness and does not necessarily favor the LCV. Indeed, if we hold onto understanding experiential states as functional states, it might be that these functional states are achieved in different beings in a way that does not rely on such cognitive machinery. To see this, think of two furniture factories both manufacturing wooden furniture with one factory employing human labor to construct, shape and assemble the furniture and one employing machine labor construct, shape and assemble the furniture; both factories at the end of the day will make the same furniture (in perhaps a more or less efficient way) but the *way* they arrive at that end result might be vastly different.

²³With the information being stored in working memory (Dehaene, 2014).

²⁴More specifically, Dehaene et al. identify four, reliable physiological markers which index when an individual is experiencing a conscious percept. They are: 1) "...an intense neuronal activation that leads to a sudden ignition of parietal and prefrontal circuits..."; 2) "...conscious access is accompanied by a solve wave called a P3 wave, which emerges as late as one-third of a second after the stimulus..."; 3) "...conscious ignition also triggers a late and sudden burst of high-frequency oscillations..."; 4) "...many regions exchange bidirectional and synchronized messages over long distances in the cortex, thus forming a global brain web" (Dehaene, 2014). It is worth pointing out that this research is still in its early phases and, as Dehaene notes, one or more of the previously mentioned markers might be epiphenomenon as the causal links have been difficult to assess (ibid). That said, the four markers have shown promise as the functional correlates of consciousness and should be taken seriously.

²⁵The standard homuncular disclaimer is required here: a briefing does not require tiny individual in the brain to brief the rest; all we need are a specialized set of processors that can manipulate information in a way that other processors cannot.

For the *LCV* to be a compelling position, some further argument is needed. Fortunately, we can articulate such an argument on behalf of the *LCV* (which I refer to as the argument from non-necessity) and we can do this by looking at the ability of the *unconscious* to process information. Dehaene articulates this ability by showing how, in experimental trials, the brains of participants reach advanced stages in processing information all *without* the information ever reaching awareness (Dehaene, 2014). The amount of information that can be processed unconsciously is vast; as Dehaene puts it:

“A great variety of cognitive operations, from perception to language understanding, decision, action, evaluation and inhibition can unfold, at least partially, in a subliminal mode. Below the conscious stage, myriad unconscious processors, operating in parallel, constantly strive to extract the most detailed and complete interpretation of our environment” (Ibid.).

Much of what might be thought of as paradigmatic conscious processes appears to be (at least partially) handled at an unconscious level and what the *GNW* *does* is to make an information state conscious and, in doing so, provides the organism with certain functional possibilities that it did not previously have. What an information state being conscious allows for is a unique way of processing that information. Very roughly put, that unique way grants the organism the ability to follow a multi-step, rational strategy that works through a problem by making it available for high-level decision making; to be conscious *is* to fulfill the functional role (Dehaene, 2014). On the *GNW*, conscious states are information states that are “broadcast” to different processors in the brain. Once broadcast, the information can, via unconscious processors, be worked on simultaneously with the results being stored in working memory. This cycle repeats itself in a serialized manner, allowing for routinized steps to be strung together and the information to be slowly and methodically processed.

By providing a functional role for conscious states and thereby illuminating the abilities of the unconscious, we get an argument for the *LCV*; roughly, experiential states are a relatively late evolutionary adaptation because earlier evolved organisms simply did not *need* to be conscious in order to engage in adaptive behavior. Of course, this is not to say that it is *logically impossible* that relatively simple biological organisms have experiential states; it is merely that they would be superfluous. It would appear that everything that *needs* to be processed in order to achieve the goals of certain, less complex organisms, can be achieved without there being anything that it is like to *be* that organism. This is what I refer to as the argument from non-necessity and it would roughly go as follows:

- 1) from the *GNW*, we can deduce what functional role experiential states play and the adaptive advantage it confers (i.e. the function of making a certain set of collated information available for high-level decision making which allows for an expanded set of behaviors that not only allows for organisms to act on the immediate circumstances but also to plan future behavior as well);
- 2) this role does not need to be fulfilled in evolutionarily older species because their behavior can be accounted for by unconscious processing;

- 3) if behavior can be accounted for with unconscious processing, then we ought not to attribute consciousness to the biological organism because, to do so, would beg the question against the function of such experiential states
- 4) the only creatures that process information in the way characterized by the GNW are more recently evolved biological organisms;
- 5) therefore, the LCV is correct and the TFV is incorrect

With the LCV and an argument for the position stated, I will move onto providing a response from Godfrey-Smith, elucidating TFV and providing two arguments for the position.

III

Godfrey-Smith and others who believe that experiential states are evolutionary old adaptations do not necessarily think that the GNW is flawed; rather, believe GNW does not provide support for the LCV, *even* with the abilities of the unconscious as discussed above. To see why, we need to consider what the GNW theory attempts to show; according to Godfrey-Smith, the GNW is best thought of as an account of *consciousness* and more specifically, human consciousness. Godfrey-Smith points out that while the notion of “consciousness” is often used interchangeably with the notions of “subjectivity”, “subjective experience” or “experiential state”, they are not strictly speaking, coextensive. He understands the notion of consciousness as a more evolved, *complex* version of subjective experience in a similar way to how a human eye is a more complex, evolved version of a sensory system that can transduce electromagnetic radiation into nerve signals (Godfrey-Smith, 2014). If this is correct, then when individuals argue that the functional role that experiential properties associated with consciousness does not need to be postulated in evolutionarily older species, they are at most providing an argument against *consciousness* in evolutionarily older creatures and not subjective experience.

More positively stated, the TFV holds that, while the GNW might be an accurate account of experiential states in humans and other mammals, there also exists experiential states that are not dependent on the sophisticated cognitive machinery described by the GNW. Supporters of the TFV hold that the experiential states described by the GNW were “reshaped” by the added cognitive mechanisms found in humans and other more recently evolved mammals (Godfrey-Smith, 2014). Godfrey-Smith’s argument for the TFV is, as best as I can reconstruct, two-fold and is comprised of what I will call: 1) the argument against special pleading and 2) the argument from logic of evolutionary explanation. I begin with the argument against special pleading.

Godfrey-Smith begins this argument by highlighting what he calls an “old form” of subjective experience that “...appear as *intrusions* into more organized and complex mental processes” (ibid). Reasonably, one might want some elucidation of what these evolutionarily older experiential states might look like and Godfrey-Smith is ready with an account. Before I begin, it needs to be acknowledged that articulation of what these earlier states might feel like *to* the evolutionarily older

creatures will, by dint of the fact that we are enmeshed in our more sophisticated, primate-centric version of these states, difficult. Indeed, as Godfrey-Smith has noted, it may be the case that we simply do not currently have the descriptive tools necessary to articulate these states (Godfrey-Smith, XXXX). That being said, we might still be able to gesture in their general direction by looking at Godfrey-Smith's discussion of "primordial emotions". As Godfrey-Smith's discussion of "primordial emotions" builds off work from Derek Denton, I will start with Denton's work.

Denton in, *The role of primordial emotions in the evolutionary origin of consciousness*, articulates what he calls "primordial emotions" as a way to understand these early forms of subjective experience. To begin, Denton defines "primordial emotions"²⁶ as "...the subjective element of the instinctive behavioural patterns" (Denton et. al., 2009). These include, "...thirst, hunger for air, hunger for food, pain, hunger for specific minerals, sexual arousal and orgasm, sensations accompanying impediment of visceral function..." and other, homeostatic-centered, emotional states (ibid). For Denton, these PEs are characterized by two aspects: 1) a sensation that, "when severe may be imperious", 2) a "compelling intention" that comes about as a result of the sensation (Denton et. al., 2009). I will proceed to unpack these two criteria.

First, PEs are understood as possibly being "imperious" in the sense that the PE can range from being slightly intrusive in the stream of consciousness to having "total occupancy" and the ability to almost completely dictate the future behavior of the organism (ibid). Pain can easily be seen to exhibit this range of influence with a dull ache of an old bruise slightly intruding in the stream of consciousness and the sharp, overriding pain of a burning limb completely dominating conscious processing and motor behavior. Second, PEs are understood to produce a "compelling intention" insofar as they produce a highly appealing, specific response to the content of the emotion. An easy example is the creation of a highly appealing intention to acquire and digest something edible in response to the PE of hunger.

With these PEs in hand, Godfrey-Smith argues for their having a felt quality to them by asking the following question: Does it seem plausible that, just because certain evolutionarily older organisms do not have the complex cognitive machinery capable of producing representations, their PEs, unlike our own, have no felt quality to them? He does not think so and, as a result, we ought to believe that they do have the felt quality in evolutionarily older creatures that have them like they do in us (Godfrey-Smith, 2018). Baldly stated, the argument against special pleading conceals the true strength of the objection and benefits from a more programmatic articulation:

- 1) In our own human subjective experience, we can plausibly identify old forms of subjectivity (what Derek Denton refers to as, "Primordial Emotions") that are correlated with certain behaviors.
- 2) Even in biological organisms that *do not* have the sophisticated cognitive machinery that appear necessary for conscious states per the GNW, we can still identify certain patterns of behavior that would be indicative of those older, subjective states in humans.

²⁶Referred to as "PE" from here on out.

- 3) If the behavioral patterns associated with the older forms of subjective experience have a certain feel to them in us, then it is plausible to think that they also have a certain feel to them in creatures that do not have such sophisticated cognitive machinery necessary for the GNW²⁷.
- 4) To argue that evolutionarily older creatures do not have such states even though they exhibit the behavior indicative of such a state would be commit a special pleading fallacy as an exception is being made without justification.
- 5) Therefore, it is plausible to think that subjective states predate the GNW and that the TFV is correct and the LCV incorrect.

With the argument against special pleading stated, I will now move onto stating the argument from the logic of evolutionary explanation.

A second reason we might adopt to the TFV is because it appears to better cohere with the logic of evolutionary explanations and to see this, we need a bit of background on how evolutionary explanations are supposed to work. One ubiquitous aspect of evolutionary adaptations is its gradual nature. With no agential management, evolutionary change is a story of small, incremental changes that are either conducive, not relevant²⁸, or harmful to the propagation of a biological organism with a certain genetic mutation²⁹. Those changes which actively help a creature to reproductive

²⁷On pg. 93 of *Other Minds*, Godfrey-Smith elaborates slightly and, as I was unclear how heavily he wanted to lean on this point, I do not include it as one of the premises, but it is certainly worth keeping in mind for this argument. He says, "Instead it seems plausible that an animal might feel pain or thirst without having an "inner model" of the world, or sophisticated forms of memory." This might be taken positively as stating that PEs are so simple that they do not necessitate an "inner model" of the world" because they can plausibly be conceived outside of such models or that PEs are so simple that they do not necessitate sophisticated forms of memory because they can plausibly be conceived of outside of such sophisticated forms of memory. I do not include this in the above stated premises because I was not sure how much Godfrey-Smith wanted to rest his claims on (at least this reading of him) conceivability claims as such arguments are notoriously controversial in the philosophy of mind (cf. Chalmers's conceivability argument against reductive theories of consciousness). I instead read Godfrey-Smith as making a plausibility claim based highlighting a special pleading fallacy in moving from human experience being dependent on patterns of behavior and then denying evolutionarily older creatures, who exhibit such behavior, having those same subjective states.

²⁸The idea that most changes are neutral relative to reproductive fitness has a bit of a rocky history; previously a controversial position and championed by Motoo Kimura, it went from a contested position to being much more accepted (Kimura, 1991). Recently, there has been some pushback with some of the stronger claims of the so-called "neutral theory" (or the more recent refinement, the "nearly neutral theory") with disagreement over what percentage of the genome can be attribute to neutral selection (Callier & Quanta Magazine). All of this is to say that individuals who tread this territory should be careful as some of the specifics are still debated and refinements are ongoing.

²⁹There is a major qualification that needs to be stated; the notion of gradualist change is, baldly stated, unclear due to an ambiguity related to the proper "unit" of evolutionary change. Intuitively speaking, the unit of evolutionary change might be thought to be individual genes but this is not necessarily the case. As Watson et. al. put it, "...in an evolutionary transition, all three of these component mechanisms (i.e., variation, selection and inheritance) are transformed or re-created at a higher level of organisation. This results in a Darwinian machine that operates via heritable variation in reproductive success at a higher level of biological organisation. There are many complex issues involved, and the details are different in different types of transition. Nonetheless, we discuss how the basic positive feedback principle applies in this context, *i.e. evolutionary units that reproduce together (at the same time or*

success will be rewarded with increased representation in the gene pool and the cementing benefit of having further mutations building off the earlier, successful mutations.

As Godfrey-Smith points out, if we want to understand the *evolution* of experiential states, we need to appreciate and find room for this gradualist aspect (Godfrey-Smith, 2018). One of the issues he points out with the LCV is how it appears to eschew this gradualist aspect and instead has experiential states as either present (when the four neurophysiological markers are met) or not (ibid). This dichotomous feature of the LCV seems antithetical to spirit of an evolutionary explanation and, as a result, should give us pause when considering whether we ought to adopt the view as an evolutionary account of experiential states.

Godfrey-Smith instead adopts the TFV where experiential states are an early evolutionary adaptation that have been modified over time. On the TFV, the GNW is understood to characterize a highly sophisticated *version* of experiential states that, again over time, have evolved to take advantage of the complex machinery present in more evolved organisms. Thus, the TFV accommodates the gradualist aspect of evolutionary explanation while *also* accommodating the data from Dehaene and others. More programmatically stated, the argument from the logic of evolutionary explanations is as follows:

- 1) If experiential states are to be explained as an evolutionary adaptation, then it ought to follow the logic of evolutionary explanation.
- 2) One necessary aspect of evolutionary explanations is the iterative process which charts the gradual change of some feature over time.
- 3) The LCV as described by the GNW does not appear to account for this iterative process.
- 4) The TFV based off PEs does account for this iterative process.
- 5) Thus, as TFV follows the logic of evolutionary explanation and it ought to be preferred to the LCV.

With both positions articulated and arguments for them laid out, I will now move onto arguing for why I think the LCV is still in a stronger position than the TFV and why one ought to prefer it. I will later, in section six, respond to the two arguments levelled against the LCV.

IV

So, with the LCV presented and the TFV responding to the argument for the LCV and two further arguments articulated in its favor, one might reasonably wonder what hope is left for the LCV. Has it not been shown to be myopic, antithetical to

under the same conditions) become 'wired' together into a new evolutionary unit, and how connectionist learning principles help us understand the consequence of this feedback in larger systems." (Watson et. al., 2015; emphasis mine). For this paper, I do not spend much time working through this complex topic because it is not directly relevant to the paper but individuals with further interest in the topic ought to keep this complexity in mind.

the spirit of evolutionary explanation and built on a rotting, fallacy-ridden core? Maybe, but I am not currently convinced and my main reason for this is that I believe that the TFV, as it has been presented by Godfrey-Smith, is does not fulfill a theoretical requirement of evolutionary accounts. To see why, we need to revisit the logic of evolutionary explanation but from a different angle; this time, focusing on its adaptive aspects. In doing this, I will be assuming the following, what I call, explanatory principle of evolutionary accounts³⁰. I formulate it as follows:

EPEA: The explanatory³¹ power of an evolutionary account primarily lies in its' explication of the adaptive advantage that the evolved function would confer on the organism(s) which possess the function. In articulating this, the account will further articulate what having such a function will allow the organism to do that it otherwise was not able to do (or at the very least, not as efficiently) without the adaptation.

To see how the EPEA is relevant to the argument I am going to put forward, we need to step back to remind ourselves of the larger picture: the main interest of this paper to articulate at what point in evolutionary history subjective states came onto the scene. To begin answering this question, I previously articulated three assumptions:

- 1) I make an assumption about what experiential states are: I will consider the experiential states of an organism as functional states (i.e., they affect the behavior of the organism) and the way they end up affecting the functioning of the organism is by modulating the way in which organism processes information.
 - 1a. I assume that experiential states are either supervenient on or identical to some information processing state of an organism.
- 2) Second, I assume the causal closure of the physical.
- 3) Third, I make an assumption about why we have these states: I assume that having experiential states provides an adaptive advantage and that advantage is a byproduct of the increased repertoire of behavior that the organism can either actively engage in or plan future behavior by simulating those possible behaviors internally.

If these assumptions are accepted, we can attempt to answer our original question by modifying it to the following: when ought we to claim that some information processing state of a biological organism has experiential properties? To answer *this* question, we might ask the following: how might we go about

³⁰Here on out referred to as the EPEA. I do not claim this to be a particularly groundbreaking principle; in fact, it seems rather innocuous, but I think codifying it can be helpful. As I take this principle to be relatively uncontroversial, I do not argue for it here.

³¹This is specifically meant to highlight the explanatory nature of *adaptive* evolutionary accounts. In specifying this, I leave out evolutionary accounts which explain the existence of some trait that is *not* adaptive (at least at the time at which it evolved; the trait might *later* confer an advantage). Stephen Jay Gould coined the term "spandrel" for such a trait and that is what I have in mind when I mention non-adaptive traits (Dehaene, 2014).

deciding whether we are justified in claiming that some information processing state of a biological organism has experiential properties?

There are at least two ways to answer this question: 1) we can directly experience³² the experiential properties and thus be justified in claiming that the organism has experiential properties. This is largely believed to be how we can be so certain of our *own* experiential states and the experiential properties contained therein and, for the most part, this internal recognition³³ is not up for dispute. That said, such a method, while perhaps high in certainty, is limited in scope with solipsistic consequences lurking around the corner. Recognizing this, we might understandably want to expand our scope and we might do so in the following way: 2) to decide whether we are justified in claiming that some information processing state of a biological organism is an experiential state, we can attempt to figure out whether the behavior of that organism is indicative of a unique style of information processing that occurs due to the presence of experiential properties.

Because of its reliance on the GNW, the advocates of the LCV can articulate the adaptive advantage that experiential states (via the modulation of information processing) provide and, thanks to the four neurophysiological markers identified by Dehaene et al., have empirical evidence by which we might test the presence of such an adaptation. Ultimately, the GNW allows for a *global processing* function; a set of information held in working memory is made available for access by different information processors in the brain which allows for the organism to *act* (or *plan*³⁴ to act) in ways that it could not previously act (or plan to act). Why could it not previously act that way? The answer from the GNW appears to be (at least) twofold: 1) it did not have *access* to the information which would allow for the high-level decision makers in the brain to make certain “choices” that would lead to those behaviors and; 2) the ability to methodically work through a multi-step procedure and deploy top-down control that allows the system to engage in complex action was not possible because the functional architecture that would allow for such processing was not in place (Dehaene, 2014).

How the advocates of the *TFV* would respond to such a question is unclear and the lack of an account provides the basis for, what I call, the argument from missing functions. In a way, the argument from non-necessity that was raised earlier attempts to press the *TFV* on just this point by asking why a certain organism, which does not fulfill the requirements of the GNW, ought to be granted experiential properties

³²Admittedly there is an implicit assumption concerning our epistemic access to our own mental states at play here in which we can directly “introspect” and pick out such properties. I should say that, while such an account is very intuitive, I am highly skeptical of such access but, for the sake of this paper, I do not go into this here. Doing so would needlessly complicate the issue at hand because my disagreement in this paper does not primarily lie in those assumptions.

³³Dependent on how we define the experiential properties that we are supposed to “find” when turning inward.

³⁴Assuming a rather weak sense of “plan”, the organism might *plan* action based off of the content of the experience. As we are considering rather simplistic organisms and long-term planning is associated with complex cognition (weighing the relative value of different actions, attempting to organize those actions in a way that is most efficacious for the accomplishing of all the actions versus organizing those actions in a way that allows for *some* of them to be definitely completed, etc.), the type of “planning” I have in mind is prioritizing the grooming and protection of a broken limb as opposed to a more superficial injury.

for the functions it achieves. Earlier in the paper, I identified an issue with this line of argument by noting how it was assuming that the functional role which the experiential properties played in organisms characterized by the GNW was the *only* functional role that might be plausibly associated with experiential properties. Contrary to this assumption, I noted how advocates of the TFV can push back by saying that perhaps experiential properties in *conscious* creatures plays such a functional role, but subjectivity is not identical with consciousness.

But if we eschew the functional role of experiential properties as discussed by the GNW because they are, at most, describing consciousness and not subjectivity itself, how are we to understand the functional role that experiential properties play in organisms described by the TFV? Another way of asking this is, on the TFV, what does an information processing state having experiential properties as opposed to it *not* having experiential properties allow the organism to *do*? Assuming it is not a neutral adaptation, then its propagation is directly due to the processing that it opens up for the organism; what might that look like on the TFV? To shed some light on this question, we can begin by looking to four functional preconditions for the presence of subjective states that Godfrey-Smith highlights. With them in hand, we might attempt to deduce what functional role subjective states might hold for an advocate of the TFV.

V

Godfrey-Smith's four preconditions³⁵ are: 1) an ability to sense and act; 2) a unique "looping" connection between sensing and acting; 3) some rudimentary form of memory; and 4) the ability to distinguish between self and other (Godfrey-Smith 79-84; Godfrey-Smith, Godfrey-Smith, personal correspondence, June 17th 2019)³⁶. The first thing to notice is that all of these tasks are information tasks; all relate to either what type of information must be present (related to sensing and acting), the relationship between the information sets ("looping"), or how the information is

³⁵From my personal correspondence with Godfrey-Smith on June 17th, 2019, I learned that he does not see the four preconditions as necessary or sufficient conditions. As he made clear, he understands the experiential states as a matter of degree and less a matter of presence/absence. Further, he notes that this is *not* to switch the "light switch" model with a "dimmer switch" model as experiential states are likely multiple gradients that play a role in determining the nature of experiential states. For this paper, I treat the four preconditions as a steppingstone from which we might be able to derive the function of experiential states, but this may change with further work from Godfrey-Smith. One worry that might be raised is the following: if we change the preconditions, wouldn't the function also change? I do not think this is necessarily correct, but I do not go into that here.

³⁶One worry is that the concepts expressed in the four criteria (specifically (1), (3) and (4)) *presuppose* the existence of experiential states (in the sense that you cannot have "sensing" without there being an experiential state of something *being sensed*) and thus cannot be one of the ingredients for the existence of an experiential state. While this is an understandable worry, it trades on an understanding of the terms *as* they are applied to experiential states and that is exactly what I am trying to avoid. "Sensing" here is supposed to be understood as the information processing aspect of a neurological state *without* the experiential properties. An analogy of this would be the aspects of a computer that process the information which are identified with (the application I am using now in fact) Microsoft Word.

registered (memory; distinguishing between self and other). This core notion, that the adaptation which confers the evolutionary advantage is an evolution in information processing is important and will become relevant in my response. With that said, I will go through each of the four aforementioned criteria and explain their importance in articulating the evolution of experiential states.

First, the ability to sense and act. If experiential states are to confer an evolutionary advantage, they need to benefit the organism which has the state. The primary benefit conferred from having an evolutionary advantage over another organism is the ability to more successfully³⁷ respond to and exploit the current/future environment of the organism. One way to achieve this benefit is to evolve sensory systems. These systems receive information from the environment (i.e. by having sensors which are responsive to chemical substances) and, with that information, adjusting activity in response. If experiential states are to provide an advantage for the organisms which have it, it will in part be an ability to better understand the environment³⁸ as one of the most important ways in which we gather information about the environment is through our senses³⁹. This information, to confer an advantage, will impact the immediate/future actions of the organisms. If the above is correct, then the category of organisms which would actually receive an adaptive advantage from having experiential states would be organisms which have some type of sensory system and can act based off the information received.

Second, we can attempt to winnow the field of possible organisms even further by articulating a certain connection between sensing and acting; a looping connection (Godfrey-Smith 79). This connection can be understood as a feedback loop in which information which is fed from the sensory systems to motor systems and, from the motor systems, back to the sensory systems⁴⁰. In this way, what we sense not only effects what action we take next but also effects what we will sense next. Godfrey-Smith highlights the importance of this by discussing tactile vision substitute systems (TVSS) for the blind (Godfrey-Smith, 2018). In these systems, there is a video camera attached to the individual which record images and transduce the information into either vibrations or electrical stimulations which are felt on the skin (ibid). After taking some time to acclimatize, the wearer starts to *experience* certain

³⁷“Success” is to be understood as a move towards an optimal prolonging of life so that a certain set of genes might be propagated into future generations.

³⁸This applies to both the internal and external environment of the organism.

³⁹This is *not* to say that experiential states are to be identified with data from the senses which, as it were, provide a bottom-up construction of the content of the state. Recent work in predictive coding seems to indicate that experiential states are largely constructed as top-down interpretations with bottom-up information from the senses confirming/disconfirming the top-down interpretation (Clark, 2013).

⁴⁰The neurological disorder known as “locked-in syndrome” presents a problem for this criterion. Individuals who have this condition have experiential states but cannot express the presence of such states because all voluntary muscle control (besides the movement of the eyes) is impaired (Smith). I think this poses a large, difficult problem for the criterion but there are a quick couple responses here: first, one might appeal to the movement of the eyes as a motor system that is changing the sensory information of the individual or; two, one might appeal to the developed capacity which was then impaired as somehow sufficient for experiential states to occur. I do not claim to have solved this problem here and will set it aside for the paper but it is worth pointing out as a possible counter-example to the criteria.

objects as located in space and not merely as a pattern of stimulation. This result, while fascinating on its own, is complicated by the fact that the experiential state *only* occurs when the blind individual is able to control the movement of the camera and modulate the incoming stream of information; thus, establishing a connection between the presence of an experiential state and the ability to modulate the incoming information in response to some prior information state (ibid).

Third, the organism must possess, at least, some kind of rudimentary memory system in order to perform non-reflexive, novel behaviors. The need for such a capacity is seen in the fact that reflexive behavior can largely be performed unconsciously; such actions like a bacterium moving towards or away from some environmental stimuli can be understood as a reflex response⁴¹ (Dehaene, 2014; Godfrey-Smith, 2018). Recall that one of the functions of an experiential state is to allow an organism to act in a way that could not be accomplished in the absence of the experiential states. Because of this, some kind of rudimentary memory system needs to be present for experiential states to provide an adaptive advantage in allowing the organism to engage in novel behaviors.

Fourth, the organism must have the ability to distinguish between “self” and “other”⁴². One central aspect of experiential states is that it appears to belong to the individual experiencing the state⁴³; the state occurs “in me” and not “in you”. The roots of the more robust ability to differentiate between “self” and “other” can be found in organisms who produce “efference copies” of motor signals to the sensory systems (Godfrey-Smith, 2018). The reason for this is that, in acting, an organism changes the sensory information that is being registered by the organism (with some of that incoming sensory information being related to the changed state of the organism) and it is useful to distinguish information that was caused by the organism and information that is the result of some action. The efference copy allows the organism to make this distinction (and counteract certain incoming information to effectively negate its influence in creating a representation⁴⁴ of the external environment) and more efficiently modulate its’ behavior moving forward. In doing this, we get the beginnings of a distinction between the “self” (that which “acts” in such a way as to modulate the incoming flow of information) and “other” (that which is external to the self). With one possible set of functional preconditions for the possession of experiential states established, we can now move onto articulating a certain function that might allow these to be achieved.

⁴¹I critique this notion of a “reflex response” later in section six of this paper.

⁴²Again, this is to be understood in both a rudimentary sense and in terms of information processing; the organism ought to possess the ability for differential registration of information.

⁴³This assumption might also be challenged. Glenn Carruthers in his 2006 paper, *A model of the synchronic self*, discusses neurological conditions in which individuals feel as if certain thoughts are not their own while still occurring “in the mind” of some individual. While, strictly speaking, the Carruthers example is not *directly* related because it focuses on “thought insertion” and not the insertion of experiential states (the difference being that the latter has experiential properties where the former does not) there might be some room to maneuver here and one ought to tread carefully before making a priori claims about experiential states belonging to the individual whose mind they occur in. Further, one might argue that thoughts themselves have a kind of phenomenology (cf. Montague 2010); for this paper, I will consider them as distinct.

⁴⁴“Representation” is being used in a very loose, non-technical sense here.

We can summarize the above stated four preconditions in the following way: the organism must be able to engage in a unique kind of information processing in which some set of information, which is registered as belonging “to” the organism, becomes part of a “looping” mechanism related to sensing/acting and this updating set of information is held in some kind of rudimentary memory system. From all of this, we might deduce the following, what I call, “orienting function”⁴⁵: subjectivity plays the functional role of orienting the cognitive processing of the organism in service of some goal by taking some information set, registering it as belonging “to” the organism, making it part of a sensing/acting “looping” mechanism that is continuously updated and held in a rudimentary memory system⁴⁶.

Denton’s PEs can be seen as states that employ the OF and I will go through each of the four criteria to show how this is the case. First, it seems as if it is only in creatures with the ability to sense and act that PEs arise as, through their compelling intention, they motivate a creature to behave in certain ways and based off sensory information, they either continue or discontinue the domination of the stream of consciousness. Second, there is a connection between such behavior and the sensory information insofar as both are necessary to get the organism to achieve a certain goal. Third, the PEs need to be held in memory in order to get the organism to, at the very least, achieve some kind of goal-directed behavior. Fourth, the PEs, in order to be effective, must be registered as belonging *to* the organism as the information coming in will constantly be updated and checked against the actions of the organism as coordinated movement and efficient action is a key requirement. If we assume that meeting the aforementioned four preconditions are sufficient for a state having experiential properties and an organism having subjectivity, then it seems as if PEs would be a prime candidate for such a state and the TFV might yet be saved.

But, as always, there are worries and I will proceed to articulate one that needs to be addressed before the OF can be seen as the functional adaptation that would allow the TFV to adhere to the EPEA. Baldly stated, it seems that the functions that characterize the OF are capable of occurring⁴⁷ in organisms in whom it is doubtful that any subjectivity arises. An example of this can be seen in Rose et. al.’s piece, *Can fish really feel pain?* when they describe the activity of decerebrated rats⁴⁸. In the article, Rose et. al. describe the behavior of the rats and the behavior seems to fit all the criteria for the orienting function. The behavior includes them reacting “...strongly to the insertion of a feeding tube, struggling, pushing at it with the forepaws and vocalizing” (Rose, 2012). From this small amount of behavior, we seem to get activity that falls right in line with the PEs previously mentioned and fit the bill for the OF. In the rats, there is certainly the ability to sense and act, that sensing and acting is part of a looping function, the rats certainly have a rudimentary

⁴⁵I will refer to this as the “OF” for short.

⁴⁶Admittedly, it is not clear what being “held” in a rudimentary memory system would mean if it is not some kind of working memory. I do not nitpick this notion for this paper, but it is a difficulty worth keeping in mind (kept, perhaps, in working memory).

⁴⁷It might be objected that non-reflexive behavior is not possible here and, as such, the functions of the OF cannot be accomplished without experiential properties present; I do not respond to this at this point because the first part of the next section responds directly to it.

⁴⁸By decerebrated, Rose et. al. point out that it is not just the cerebrum that is removed but “...all of the brain above the midbrain including the diencephalon, cortex and subcortical forebrain” (2012).

memory system and a basic notion of self yet the idea that, in the absence of all brain structure above the midbrain, there still exists some semblance of subjectivity is, at the very least, doubtful. In response, one might dig their heels in and contend that, in this case, even in the absence of such structures, some form of subjectivity exists as the rat would not be able to orient its functioning toward the needle without it. One can plausibly respond this way but it seems that the burden of proof is on the advocate of the TFV to explain why one needs to posit experiential states in order to explain the function of the (relatively) basic activity of the midbrain (mostly associated with motor activity) and hindbrain (mostly associated with autonomic functions). Until the advocate of the TFV provides some answer, it seems as if we do not *need* to posit experiential states in order to get the OF on board; all of the information processing that needs to happen appears possible, as it were, “in the dark”. That said, I will now move onto defending the LCV from the argument against special pleading and the argument from the logic of evolutionary explanation and then conclude.

VI

I begin with the argument against special pleading. One of the cornerstones for the argument is its claim that we can identify, within ourselves, certain *old* forms of subjectivity that are associated with specific behavioral patterns. To understand this, we need to get clear on why Godfrey-Smith claims certain forms of subjectivity to be *old* forms of subjectivity. In full, he says: “The best argument I can offer for this alternative view [TFV] is based on the role in our lives of what seem like old forms of subjective experience that appear as *intrusions* into more organized and complex mental processes” (Godfrey-Smith, 2018). These “old forms of subjective experience” are understood as the PEs that I discussed above. To see where the argument is flawed, I will recount the argument from special pleading:

- 1) In our own human subjective experience, we can plausibly identify old forms of subjectivity (what Derek Denton refers to as, “Primordial Emotions”) that are correlated with certain behaviors.
- 2) Even in biological organisms that *do not* have the sophisticated cognitive machinery that appear necessary for conscious states, per the GNW, we can still identify certain patterns of behavior that would be indicative of those older, subjective states in humans.
- 3) If the behavioral patterns associated with the older forms of subjective experience have a certain feel to them in us, then it is plausible to think that they also have a certain feel to them in creatures that do not have such sophisticated cognitive machinery necessary for the GNW.
- 4) To argue that evolutionarily older creatures do not have such states even though they exhibit the behavior indicative of such a state would be commit a special pleading fallacy as an exception is being made without justification.
- 5) Therefore, it is plausible to think that subjective states predate the GNW and that the TFV is correct and the LCV incorrect.

To see the issues with this argument, we need to get clearer on the third premise; more specifically, the idea that a similarity in behavioral patterns makes it plausible to believe that the organism which exhibits the behavior is experiencing something sufficiently similar⁴⁹ to the experience that we have when we exhibit that behavior. For this argument to work, one has to assume that the presence of such behavioral patterns is *sufficient* for the presence of the experiential states and not just necessary. But it is not at all clear why the LCV ought to accept such an assumption as one of the key aspects of their claim is that experience is to be correlated with certain information broadcasting states and not necessarily patterns of behavior.

In response to this claim, the defender of the TFV might retort as follows: it is not that behavioral patterns simply *correlate* with experiential states; it is that certain patterns of behavior (those associated with PEs) evoke in humans an experiential state and that experience *part and parcel* of what it is to be able to engage in those behavioral patterns. There seems to be some textual evidence for his interpretation; in *Other Minds*, Godfrey-Smith notes how it is not necessarily the case that all instances of certain behavioral patterns that overlap with human behavior are indicative of experiential states, rather it is a certain *type* of behavior: non-reflexive behavior⁵⁰ (Godfrey-Smith, 2018). As an example, Godfrey-Smith alludes to the activity of zebra fish whose behavior appears too sophisticated to be understood as reflexive but their (relatively) basic biology seems to disqualify them from the LCV. Put more precisely, the advocates of the TFV might respond by saying that non-reflexive behavior is a sufficient condition for the positing of experiential states and as certain PEs exhibit this non-reflexive behavior, we should posit experiential states in creatures that predate the GNW.

To see what motivates Godfrey-Smith's point, it helps to understand the study he is referencing. The study is roughly as follows: zebrafish were initially exposed to two different environments with the researchers determining which was more preferred and which was less preferred. After this, they were injected with a chemical that is thought to cause pain and were allowed to choose which, between the two environments, they wished to inhabit. In the less preferred environment, experimenters dissolved a painkiller and did not change the other, more preferred environment. The zebrafish, after being injected with the chemical, tended to occupy the less preferred environment, presumably, *because* of the effects of the painkiller on the experiential state of the fish (Godfrey-Smith, 2018). According to Godfrey-Smith,

⁴⁹I say "sufficiently similar" because it seems implausible for there to be a complete or even near identical experience as experience, at least in part, is influenced by the contingent circumstances that an individual is in. For example, consider the difference between experiences where an individual is in pain but has the potential for relief as opposed to the circumstance where the individual is in pain and that pain is anticipated to continue on. At this point, one might dig in their heels and disagree by discussing a circumstance in which an individual is *fully consumed* by the pain experience with second-order thoughts about the pain being available. While I am suspicious of such a move, I can see how one might appeal to it and do not say more about it here.

⁵⁰This is interesting as the ability to engage in non-reflexive behavior (of a certain sort; I will elaborate on this in the next paragraph) is a condition that Dehaene identifies as one of the functions of consciousness and while Godfrey-Smith seems to appeal to this function as indicative of conscious experience (cf. his explicit reference on pg. 91 of *Other Minds*), he does not go so far as to acknowledge the further idea that this complex behavior happens because of the distributed, parallel processing which the GNW allows.

evolution could not have set zebrafish up with a reflexive reaction to this situation and thus, in order to explain the behavior, we need to posit experiential states as playing some mediating role (ibid).

In response, we can note that the non-reflexive behavior that Godfrey-Smith is alluding to in the zebrafish example, is not necessarily indicative of experiential states. To see this, we can turn again to Rose et al. In their paper, they draw attention to the distinction between nociception and pain experience. Nociception is the registration of painful stimuli through nociceptors (pain registration receptors) and, *by itself*, is not pain; that is to say that an identification between the activation of nociceptive pathways and pain would be incorrect. According to Rose et. al., at most, activation of nociceptive pathways appears as a necessary but not sufficient condition for an actual pain state.

This links up with Godfrey-Smith's work insofar as Rose et. al. highlight how studies which claim that fish must be feeling pain due to the presence of non-reflexive behavior have not actually provided an adequate criterion on what "non-reflexive" behavior is supposed to mean as nociceptive response (in the absence of pain) seems to give rise to all types of behavior that might be termed, "non-reflexive".

They discuss this point by noting how individuals often claim that "non-reflexive" behavior is "more than a simple reflex" or something similar and this, by itself, seems to rely on assumptions about the nature of "simple reflex" behaviors that are themselves unmotivated (Rose 103). To see this point, we can look to the behavior of rats who, due to damage in their central nervous systems, seem to exhibit quite complex behavior that fits uncomfortably with the notion of "simple reflex" behavior.

To start, we can begin with Godfrey-Smith's discussion of rat behavior how rats, even with their spinal cords severed, still exhibit pain behavior⁵¹. Commenting on this, Godfrey-Smith notes how the pain behavior (including vocalizing, coordinated muscle movement, biting and licking) in these rats is possible outside of consciousness because evolution might plausibly have programmed the biological organism for such reflex reactions.

Rose et. al. respond directly to such a claim and show how decerebrated rats (as mentioned previously at the end of section five) can activate certain highly coordinated motor responses in an attempt to avoid noxious stimuli that appears complex but is not the product of conscious processes (Rose, 2012). Recall that Rose et. al. mention decerebrated rats that "...react strongly to the insertion of a feeding tube, struggling pushing at it with the forepaws, and vocalizing" and "...these rats react indistinguishably from a normal rat: vocalizing, attempting to bite the syringe or experimenter's hand, and licking the injection site." (Rose, 2012). In this example, Rose et. al. claim that this appears to be, on an everyday understanding of "simple reflexes", much more than a simplex reflex yet it still appears to happen unconsciously⁵². If this is correct, then it seems as if we can understand quite complex actions as occurring unconsciously and the upshot of this is one of two outcomes: 1) either Godfrey-Smith is incorrect in understanding non-reflexive behavior as requiring consciousness or, 2) we can broaden the notion of "reflexive behavior" to include not only certain narrow actions but

⁵¹Thus, cutting off the connection between the site of the bodily damage and the brain.

⁵²See the end of section five for more on this "unconscious" claim.

certain action *types*⁵³. I imagine that Godfrey-Smith would want to go with the latter but if we go with the latter, then it does not appear that zebrafish (and other creatures of this complexity) would require the attribution of conscious states as their behavior could plausibly be described without appeal to experiential states.

For example, we can imagine a fish in the painful environment with the pain causing a nocifensive response. This nocifensive response would be carried out by the fish moving from one environment to the other (less preferred) environment with the difficulty arising for explaining why the fish would move to the less preferred environment. Here, the notion of an action type is helpful; we might pick out one “action type” as a “protective” action type that is sensitive to the activation of the nocifensive pathway and, with a minimal amount of environmental information, acts to mitigate the nocifensive response. As an analgesic can work to desensitize nociceptors⁵⁴, the fish would activate a swim response with a minimal amount of environmental information until it is in the (previously) less preferred environment (Mense 4).

If what I have said above is correct and the behavioral patterns that are thought to overlap with humans can be understood as nocifensive responses that do not necessitate experiential states, then it seems as if the third premise is false and more work needs to be done to modify it, or other premises, in such a way that the argument goes through. If one wants to continue to insist that earlier evolved organism’s PEs must feel like something because it feels like something *in us*, they would be begging the question against the LCV as its’ advocates are not arguing for an identification between behavioral patterns and experience but rather between information distribution and experience. With that said, I will now move onto responding to the argument from the logic of evolutionary explanation. To begin, I restate the argument here:

- 1) If experiential states are to be explained as an evolutionary adaptation, then it ought to follow the logic of evolutionary explanation.
- 2) One necessary aspect of evolutionary explanations is the iterative process which charts the gradual change of some feature over time.
- 3) The LCV as described by the GNW does not appear to account for this iterative process.
- 4) The TFV based off PEs does account for this iterative process.
- 5) Thus, as TFV follows the logic of evolutionary explanation and it ought to be preferred to the LCV.

To see how a supporter of the LCV might respond to such an argument, it helps to home in on one of the premises and, like the argument against special pleading, the third premise is a good place to begin. The third premise notes how the LCV, as

⁵³I do not spend much time fleshing out this notion in this paper: intuitively what I have in mind is the idea is the distinction between a specific action (such as moving left or moving right) and a more general set of actions in service of a certain end (such as avoidance or protective behavior). The latter (action types), can be understood as more complex insofar as they allow for a very minimal amount of environment information to contribute to coordinated action in service of a basic goal whereas the former (specific actions) are less complex insofar as it does not take in any information about the environment into account in the formation of the output.

⁵⁴Depending on the analgesic; different drugs have different methods of activation.

described the GNW, does not appear to account for the iterative process of evolutionary adaptation and the supporter of the LCV might respond by closely analyzing *how* the LCV does not account for such a process.

From premise two, we receive an account of *what* iterative process is being picked out; it the “gradual change of some feature over time”. Taken together then, we can rephrase premise three as saying that the LCV as described by the GNW does not appear to account for the gradual change of some feature over time. But it is not at all clear that the LCV does not do this and to see how, we need to look more closely at the idea of the “gradual change of some feature over time”.

The reason, it would seem, that the LCV is thought to not account for the gradual change of some feature over time is because the LCV, due to its’ reliance on the GNW, appears to move from a state of affairs in which an animal does not have experiential states and then, due to some fortuitous mutation, to have the necessary resources with experiential states abruptly coming online. But it is not at all clear that this characterization of the GNW is necessary and the reason for this is that the argument does not consider the possibility of the gradual changes in information sharing which in turn might correlate with more or less qualitatively rich experiential states.

For example, we might consider less sophisticated forms⁵⁵ of information sharing in creatures that are earlier in evolutionary history than paradigmatic cases of the GNW and charting the gradual steps from this less sophisticated style of processing towards a more sophisticated style. Instead of adopting this viewpoint, the argument considers the existence of experiential states in a binary manner instead of a “dimmer switch” model in which the states might change over time. If we do adopt the “dimmer switch” model, it seems we can accommodate the gradualist aspect of evolutionary explanation.

In response to this response, an advocate of the TFV might note the rather ad-hoc nature of this appeal; part of what it was that made the LCV such a troubling position is the high cognitive demands that it places on certain creatures and remarking that it may come in degrees seems to introduce the idea of gradations in experiential states without justification. There are at least two responses to the ad-hoc claim. First, we may respond by noting that the idea of their being more or less information integration in a given conscious experience is not nearly as alien as it seems, and we experience these moments quite regularly.

For example, consider waking up from a deep sleep, being on some kind of drug that has dissociative effects, or even being sleep deprived: in all these instances, our everyday waking consciousness is disturbed with one or more of sensory systems playing a different role than normal. This abnormal conscious state in no way threatens the fact that it *is* a conscious state that plays a functional role in determining behavior.

Second, we can note that all of the cognitive mechanisms that the GNW relies on are *themselves* products of evolution and, as such, they come in more or less sophisticated forms. Given that, we might respond to the critique by reaffirming that the relatively sophisticated cognitive mechanisms need to be in place but the

⁵⁵Insofar as there is less distribution of information to various processors of the brain.

mechanisms themselves might be less evolved while still fulfilling a certain functional role (albeit in perhaps a less effective way).

For example, one of the key cognitive mechanisms of the GNW is working memory; we might understand working memory as more or less robust in the sense that more or less information might be stored in working memory⁵⁶ (Dehaene, 2014). In this sense, we might understand the information held there to be more or less qualitatively rich and without necessarily changing the criteria for the GNW. Understanding it in this way would allow us to see how the LCV might accommodate the gradual evolution of experiential states in a way that still remains true to the theory and is not ad-hoc. With all that said, I will finally conclude.

Conclusion

I see the upshot of this paper as a shifting of the burden of proof onto the holders of the TFV to explain the adaptive advantage that experiential states confer on those creatures that can exhibit the same behavior in the absence of consciousness. It is *not* to say that the TFV is incoherent, logically flawed or otherwise crippled; it is rather that advocates of the position need to do more to explain why one ought to buy into the view in the first place.

As I stated at the outset, this paper is not one that is attempting to provide necessary and sufficient conditions for the presence of experiential states: such an analysis appears to me to be largely an empirical matter. This paper is more concerned with orientation: should we understand experiential states as they are understood by the GNW and the functional benefit that it provides or should we look instead in a different direction (perhaps at primordial emotions and an information set dominating the cognitive operations of some organism and providing a compelling intention)? I argued for the former by showing how, while relatively demanding, the LCV is adept at providing an evolutionary explanation of experiential states as it follows the EPEA. For the advocate of the TFV to provide a compelling response, they need to take seriously the EPEA and clearly explicate why less evolved organisms warrant the attribution of experiential states.

References

- Braitenberg V (1987) *Vehicles: Experiments in Synthetic Psychology*. The MIT Press, 1987.
- Callier Viane (2018) Neutral Theory of Evolution Challenged by Evidence for DNA Selection. *Quanta Magazine*, <https://www.quantamagazine.org/neutral-theory-of-evolution-challenged-by-evidence-for-dna-selection-20181108/>.
- Calvo P (2017) What Is It Like to Be a Plant? *Journal of Consciousness Studies*, vol. 24, no. 9-10, 2017, pp. 205–227., http://www.esalq.usp.br/lepse/imgs/paginas_thumb/Whats-Is-It-Like-to-Be-a-Plant.pdf.

⁵⁶Specifically, Dehaene notes how children might be thought to have a rudimentary form of verbal working memory.

- Carruthers G (2007) A Model of the Synchronic Self. *Consciousness and Cognition*, vol. 16, no. 2, 2007, pp. 533–550., doi:10.1016/j.concog.2006.06.002.
- Clark A (2013) Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. *Behavioral and Brain Sciences*, vol. 36, no. 3, Oct. 2013, pp. 181–204., doi:10.1017/s0140525x12000477.
- Dehaene S (2014) *Consciousness and the Brain Deciphering How the Brain Codes Our Thoughts*. Viking, 2014.
- Denton D, et al. (2009) The Role of Primordial Emotions in the Evolutionary Origin of Consciousness. *Consciousness and Cognition*, vol. 18, no. 2, 2009, pp. 500–514., doi: 10.1016/j.concog.2008.06.009.
- Godfrey-Smith P (2018) *Other Minds: The Octopus and the Evolution of Intelligent Life*. William Collins, 2018.
- Herculano-Houzel S (2009) The Human Brain in Numbers: a Linearly Scaled-up Primate Brain. *Frontiers in Human Neuroscience*, vol. 3, 2009, doi:10.3389/neuro.09.031.2009.
- Huang EP (1998) Synaptic Transmission: Spillover at Central Synapses. *Current Biology*, vol. 8, no. 17, 1998, doi:10.1016/s0960-9822(98)70389-6.
- Jasanoff A (2018) *The Biological Mind: How Brain, Body, and Environment Collaborate to Make Us Who We Are*. Basic Books, 2018.
- Kimura M (1991) The Neutral Theory of Molecular Evolution: A Review of Recent Evidence. *The Japanese Journal of Genetics*, vol. 66, no. 4, 1991, pp. 367–386., doi:10.1266/jgg.66.367.
- Levine J (1983) Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, vol. 64, no. 4, 1983, pp. 354–361., doi:10.1111/j.1468-0114.1983.tb00207.x.
- Mense S (1983) Basic Neurobiologic Mechanisms of Pain and Analgesia. *The American Journal of Medicine*, vol. 75, no. 5, 1983, pp. 4–14., doi:10.1016/0002-9343(83)90226-7.
- Montague M (2010) Recent Work on Intentionality. *Analysis*, vol. 70, no. 4, 2010, pp. 765–782., doi:10.1093/analys/anq090.
- Oda S, et al. (2017) Modulation of Sensory Information Processing by a Neuroglobin in *Caenorhabditis Elegans*. *Proceedings of the National Academy of Sciences*, vol. 114, no. 23, 2017, doi:10.1073/pnas.1614596114.
- Organism (2014) *Biology*, 12 May 2014, <https://www.biology-online.org/dictionary/Organism>.
- Pulvermüller F, et al. (2014) Thinking in Circuits: toward Neurobiological Explanation in Cognitive Neuroscience. *Biological Cybernetics*, vol. 108, no. 5, 2014, pp. 573–593., doi:10.1007/s00422-014-0603-9.
- Rose JD, et al. (2012) Can Fish Really Feel Pain? *Fish and Fisheries*, vol. 15, no. 1, 2012, pp. 97–133., doi:10.1111/faf.12010.
- Smith E (2005) Locked-in Syndrome. *BMJ*, vol. 330, no. 406, 17 Feb. 2005, <https://www.bmj.com/content/330/7488/406>.
- Watson RA, et al. (2015) Evolutionary Connectionism: Algorithmic Principles Underlying the Evolution of Biological Organisation in Evo-Devo, Evo-Eco and Evolutionary Transitions. *Evolutionary Biology*, vol. 43, no. 4, Aug. 2015, pp. 553–581., doi:10.1007/s11692-015-9358-z.
- White JG, et al. (1986) The Structure of the Nervous System of the Nematode *Caenorhabditis Elegans*. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 12 Nov. 1986, <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.1986.0056>.