How to Relate Major Transitions in Life and Cognition?

Farid Saberi, 2025

Pre-Print (to be published in Biology and Philosophy)

Abstract. Recent innovative research has focused on major transitions in cognitive evolution, drawing from the existing literature on major transitions in the evolution of life. This prompts a careful examination of the distinctions and similarities between these two types of transitions. In this paper, I present four claims. First, a theoretically fruitful approach to understanding major evolutionary transitions (METs) in life is to conceptualize them as a set of objectively similar events, akin to a natural kind concept. Second, this framework allows for discussing major cognitive transitions (MCTs) while emphasizing that METs and MCTs represent two distinct subsets of possibility-expanding evolutionary events, each defined by different criteria. Third, the recent works of Barron et al. (2023) and Ginsburg and Jablonka (2019, 2021) serve as successful examples of applying a transition-oriented approach to cognitive evolution. Both provide coherent definitions of MCTs along with fine-grained explanations of these events in unique ways. Finally, drawing on the tradition of dialectical thinking, specifically the method of climbing down the ladder of abstraction, I argue that their contributions can be viewed as complementary rather than competing alternatives.

1- Introduction

In recent years, scholars have started to take a transition-oriented approach to the evolution of cognition and the brain more seriously. This approach has proven fruitful in studying the evolution of life. It highlights certain evolutionary changes that lead not only to adaptive traits but also to expanding the possibility of evolvability and the emergence of new units of selection subjected to new forms of natural selection. Inspired by the work of Maynard Smith & Szathmáry (1995) on major transitions in the history of life, many scholars have offered major transition proposals for cognitive evolution. Most proposals are focused on changes in domain-general or domain-specific cognitive capacities that open up new space for more cognitive complexity and evolvability. Different cognitive capacities that these proposals have highlighted include associative learning, recursive syntax, trial-and-error exploration, use of symbols, joint intentionality, and collaborative computation (Ginsburg & Jablonka 2010, 2021, Hauser et al. 2002, Dennett 1996, 2017, Dehaene et al. 2022, Tomasello 2014, Dor 2023).

The connection between concepts of major evolutionary transitions in life (MET) and major cognitive transitions (MCT) is not straightforward. Some authors understand MCT as a subset of MET or at least see substantial overlap between them. Others consider them similar but separate (Szathmàry 2015, Ginsburg & Jablonka 2019, 2021, Barron et al. 2023). In this paper, I argue that a useful way to understand them is to view them as two separate natural kind concepts with different criteria that identify distinct sets of events. Interpreting events in the MCT list as a subset of the MET list of METs would be conceptually confused and theoretically unfruitful. In line with the arguments of some philosophers of biology, I contend that any serious account of METs must meet certain conceptual and theoretical requirements regarding how they define and explain METs. I will discuss requirements such as a uniform definition of METs (theoretical unity of the list of METs) and fine-grained evolutionary explanations of them (Okasha 2022, Herron 2021, McShea & Simpson 2011, Calcott 2011). I draw on the homeostatic-property-cluster (HPC) theory of natural kind (Boyd 1999) to clarify the link between these two requirements.

Next, I argue that any serious account of MCT must similarly meet the requirements of theoretical unity and fine-grained explanations. Two recent proposals in cognitive evolution that have explicitly used a transition-oriented approach are the learning-based proposal of Ginsburg & Jablonka (2019, 2021) and the "computational architecture" proposal of Barron, Halina & Klein

(2023). I argue that both of these proposals meet the requirements of theoretical unity and fine-grained explanations. However, they do so in different ways. There are interesting points of convergence between their approaches. Information processing and its biological efficiency play a central role in both proposals. Their lists of MCTs can be more or less mapped onto each other. However, there are still differences. They uphold different definitions of cognition and MCTs. Also, the fine-grained explanation in Barron et al.'s proposal is based on a specified aspect of natural selection, that is, an explanation based on resource constraints and the immediate fitness value of energy efficiency. In contrast, Ginsburg and Jablonka use an extended view of natural selection and an evolutionary-developmental approach. They highlight different aspects of evolutionary processes accounting for MCTs. In terms of incorporating the evo-devo approach and the role of developmental plasticity in the evolution of cognition and the brain, Ginsburg and Jablonka's account is more developed. Finally, I discuss the possibility of considering their different definitions and explanatory strategies as supplementary. Their different definitions could be considered as grasping MCT at different levels of abstraction, and their explanations as accounting for different aspects of the same evolutionary process. In making this point, I draw on the philosophical tradition of dialectical thinking (as articulated by philosophers like Hegel and Marx), namely, the method of climbing down the ladder of abstraction and modifying definitions as the process of explanation proceeds.

To achieve these goals, in Sect. 2, I discuss the different restrictive and permissive conceptions of MET and consider the proposed requirements for an adequate definition and explanation of METs. In Sect. 3, I will evaluate what makes Ginsburg & Jablonka's and Barron et al.'s proposal unique among other approaches to cognitive evolution. Then, I compare them and their compliance with requirements such as theoretical unity and fine-grained explanations. In Sect. 4, I assess the similarities and differences between their proposals and conclude by discussing how these proposals could be further developed (Sect. 5).

2- Conceptualizing transition in life and cognition

The origin of transitional thinking in evolutionary biology can be traced back to the works of Buss (1987), Maynard Smith and Szathmàry (1995), and Michod (1999). The common insight among these works was that there are specifiable, unique events in evolutionary history, distinct from other types of normal changes and even more complex and novel evolutionary changes. Moreover, one must explain how and why they occur with the aid of explanatory principles that are articulated more specifically than merely mentioning general principles such as fitness and natural selection used for typical evolutionary changes (Okasha 2022). In other words, from a big-picture perspective of evolutionary history, there are major transitions that go beyond trivial adaptations in local populations or even the emergence of complex or novel traits in lineages. These transitions need to be identified and explained with specific criteria.

Before jumping into discussions of different ways of defining and explaining METs, some terminological considerations and clarifications are in order. Life is a multi-dimensional phenomenon. Therefore, one can look at the history of life and its major transitions from different perspectives. A key tenet of the Darwinian way of looking at the history of life is in terms of changes in biological populations or what Mayr (1982) has famously called population thinking. From this perspective, major transitions in life would most naturally appear primarily as the emergence of new evolutionary populations and individuals. Authors like Michod (1999, 2011) and Godfrey-Smith (2009) have used the term transition in this way. However, it is also possible to have, say, an Aristotelian or informational view of life. Authors like Ginsburg and Jablonka (2021, 2019) and Dennett (1996, 2017) have talked about ecological transitions, transitions in intentionality, kinds of mind, and Aristotelian teleological transitions in the mode of being (nutritive life, conscious life, rational conscious life). These different perspectives on the history of life and its major transitions can sometimes be translated into one another. For example, for Ginsburg and Jablonka, the transition to consciousness could at the same time be described as a transition in Aristotelian teleological modes of being (sensitive soul), ways of controlling information, and types of learning. However, this is not always the case. For example (as will be discussed in Sect. 2.1), the transition to language from an informational perspective does not have a direct equivalent from the perspective of transitions in Darwinian populations. Moreover, not all these perspectives and their associated list of transitions allow for the same level of fine-grained explanations. For example, the fitness values (of cooperative behavior of the units) that explain the emergence of multicellularity and eusociality are similar in kind. In contrast, the meaning of the term "information" and the evolutionary pressures influencing the origin of DNA information and linguistic information are distinct.

In this paper, I have primarily focused on the transitions in life from a Darwinian population perspective. Therefore, when I consider the relationship between MCT and MET, I understand MET in a Darwinian framework. There are two reasons for this terminological decision. Firstly, given the limited space and scope of the paper, one has to focus on one perspective and its merits while acknowledging the usefulness of other perspectives on the history of life and transitions. Secondly (and more importantly), as it will become clear during the course of this paper, Michod and Godfrey-Smith's Darwinian approach provides us with a natural kind view that allows for uniform definitions and fine-grained explanations of MET. I have argued that it is useful to apply the same natural kind approach for MCTs and consider MET and MCT as two distinct categories with no overlap. [1]

2.1. Defining MET

Major transitions were understood to be partially similar to the emergence of novel and complex traits with outsized evolutionary and ecological outcomes, like the emergence of photosynthesis and vision. They both expand the range of possibilities for the evolutionary process, affect the biosphere in outsized ways compared to other evolutionary events, and increase the evolvability of lineages, that is, the potential of a lineage to show a higher rate of diversity and novel traits. However, there has been some inconsistency in the use of the term, and some authors have argued that MET should be used in a way that possibility-expanding changes is a necessary but not sufficient condition for applying this concept. This view restricts the use of the term to events that not only expand biological possibilities and have outsized effects but also involve the formation of new higher-level units of selection from collectives of smaller evolutionary units. For example, the formation of multicellular organisms from single-cell eukaryotes (Calcott 2011, Michod 2011, Okasha 2022, Herron 2021).

---

[1] Note that my claim here is not the primacy of the Darwinian population framework, nor that every other perspective on the history of life should be translatable or find an equivalent in the Darwinian individuality list. Rather, my claim is that each list of transitions from any perspective on life (Darwinian, Aristotelian, informational) should (more or less) meet the two requirements of uniform definition and fine-grained explanations, granted that one is interested in the epistemic merits of the natural kind view. The fact that some of the transitions on two lists may or may not overlap just reflects the fact that there are different, equally legitimate perspectives to look at the evolution of life. These different perspectives might have different epistemic merits. One could be more powerful in making sense of long-term patterns in the history of life, and the other could allow for more operationalized definitions and fine-grained explanations. As will be discussed in Sect. 2, Maynard Smith & Szathmàry's (1995) original list is decomposable into two distinct lists (Darwinian individuality and informational list). The first one meets the two requirements, while the second one does not because the notion of information is too vague and broad to be operationalized in the case of all the transitions. They also do not overlap on all the items. McShea and Simpson (2011) have mentioned this problem of vagueness for different proposed transition lists, and especially Maynard Smith & Szathmàry's list. I want to especially thank the anonymous reviewer for highlighting the importance of clarifying this point.

As Okasha (2022) has argued, the reason behind this inconsistent, restrictive, and permissive use of the term in the literature goes back to Maynard Smith and Szathmàry's (1995) "two-fold characterization" of METs. They use two criteria that do not always go hand in hand; firstly, those changes that involve the emergence of higher-level evolutionary individuals from smaller hierarchically nested units. In this way, "entities that were capable of independent replication before the transition can replicate only as part of a larger whole after it" (Maynard Smith & Szathmáry 1995, 4). Secondly, those changes that involve changes in the way that developmental information (genetic and linguistic) is transmitted and stored over generations.

I refer to these two criteria as higher individuality and information criteria. According to these criteria, they proposed a list of METs with eight transitions: (1) from replicating molecules to protocells; (2) from gene to chromosome; (3) from RNA to DNA world; (4) from prokaryotes to eukaryotes; (5) from asexual clones to sexual population; (6) from single cell eukaryotes to multicellular organisms; (7) from solitary individuals to colonies; and from (8) from primate sociality to human societies with language (Maynard Smith & Szathmàry 1995, 6). As one might notice, not all the items in their list meet both of the criteria. For example, the transition from RNA to DNA or the origin of sex and human language does not involve higher individuality, and they only work with the information criterion. This led some scholars to raise concerns about "theoretical unity" and a uniform definition of METs and describe their list as a miscellaneous or arbitrary grouping (Queller 1997, McShea & Simpson 2011, O'Malley and Powell 2016, Heron 2021, Okasha 2022).

Szathmàry (2015) has defended this two-fold characterization (with some modification) as a feature and not a bug of their account.[2] He argues that although these two criteria are "conceptually independent," there is a tractable "empirical link" between them. In many cases, changes in information transmission and storing are a prerequisite for higher individuality, e.g. emergence of an epigenetic system of information transmission was necessary for the emergence of multicellularity as higher individuality. Some authors like Herron (2021) and Okasha (2022) have not found this defense convincing and have argued for a uniform definition of MET based solely on the higher individuality criterion (cf. O'Malley & Powell 2016). This way of defining constitutes METs as a coherent category so that one can make explanatory and useful generalizations based on it. The information criterion and how it relates to higher individuality in every case remain vague and underdeveloped in Maynard Smith and Szathmàry's account. This might endanger the very premises underlying the transition-oriented approach, namely, there is a set of specifiable, unique events that need specific and fine-grained explanations. For these reasons, they have proposed to take the higher individuality criterion as defining METs and hence exclude some items from the original list (like the origin of genetic code, the transition from RNA to DNA, and the origin of sex and human language). In this way, the transition in the restrictive list shows objective similarity and can be treated as a "natural kind", that is, "a set of events (or objects) that are objectively similar to each other as opposed to an arbitrary grouping" (Okasha 2022, 2).[3]

---

[2] Szathmàry (2015) excluded the transition to sex from his revised list and added Plastids. For a detailed table comparing different lists, see Herron (2021, 4).

[3] Authors favoring a natural kind definition of METs, like Herron (2021), have argued that one should only focus on the higher individuality criterion and that the evolutionary and ecological outcome of METs should not be treated as a defining criterion for METs. This approach risks rendering the definition of MET too abstract for the concrete and multidimensional reality of the evolutionary history. I will come back to this point and offer an enhancement to this view at the end of Sect. 4, where I discuss levels of abstraction and modifying definitions that incorporate different aspects of phenomena as the process of explanation proceeds.

This approach aligns with Michod’s approach that highlights transitions in individuality and treats them as a natural kind, which he defines as "a natural grouping of phenomena involving common problems and sharing common solutions" (Michod 2011, 170). This resonates with Maynard Smith and Szathmàry’s (1995) original hope that “sufficient formal similarity” among METs can enable us to infer features of others by “progress in the understanding of any one of them” (Maynard Smith & Szathmàry 1995, 23). In other words, if we have a uniform definition of MET, we can hope to find common properties across all transitions. For example, cooperation of lower-level units (Michod & Herron 2006) and division of labor (Michod 2007) are common properties across all METs in this sense. There is a co-occurrence between the property of higher individuality and the property of cooperation and division of labor. Hence, when one sees one of these properties can start to look for others and infer their presence.

METs are perceived as processes and not a group of objects, organisms, or traits. However, some argue that their objective similarity could give them a role like natural kind concepts in biological explanations and generalizations. For example, Brett Calcott (2011) has argued that the formal similarity and hence the ability to generalize from one MET to others is analogical generalization and not homological. In the case of biological traits, homologous traits are those traits in different species that are similar because of common descent. Analogous traits are similar because each species has separately evolved those traits in response to the same universal environmental conditions. In the latter, evolution has made similar “forced solutions” in the face of the universal physical constraints of the environment. Now, let us change our perspective from the similar traits of separate species to the similar properties of separate METs. All of them show cooperation and division of labor. This similarity is analogous. According to Calcott, they are separate but similar responses to the physical constraint of the environment. This view of useful explanatory generalization based on analogy further supports the idea of treating METs like natural-kind concepts.

I uphold this restrictive definition based on the higher individuality criterion. It highlights the importance of treating METs as a natural kind and allows for useful explanatory generalizations and inferences. Authors like Richard Boyd (1999) have advocated for a homeostatic-property-cluster (HPC) theory of natural kind in biology according to which a collection of objects could be considered as a natural kind if and only if they meet two criteria: (1) they share co-occurring hemostatic properties in a way that presence of some favor the presence of the others or an underlying mechanism maintain all of them together and (2) the natural kind concept play a useful role in explanatory generalizations in an inferential domain (Boyd 1999).

Granted this view of natural kind concepts, METs could be considered as a set of evolutionary events that show homeostasis properties. In the restrictive notion of METs, all of them are instances of higher individuality in biological populations constructed from smaller units in a way that smaller units are not able to replicate independently, show internal division of labor, and involve higher levels of selection on the new units. Some authors have argued that, like other natural kind terms, METs have sub-kinds depending on whether the collection of units is of common descent (fraternal) or not (egalitarian) (Queller 1997, Herron 2021, Okasha 2022). [4]

---

[4] The debate about different ways of defining MET and dividing it into sub-kinds cannot be considered a settled debate. As mentioned above, one party to this debate includes authors like Szathmàry (2015) and Ginsburg & Jablonka (2019) who think a two-fold characterization of METs (or seeing transitions simultaneously from Darwinian individuality and informational perspective) gives us a coherent category because the empirical link between the two conceptually distinct criteria is (at least in principle) tractable. Under this view, not only fraternal-egalitarian sub-kinds but also transitions in individuality and in ways of control of information are two legitimate sub-kinds of the METs. Another camp in this debate (Okasha, Herron, Michod, and Godfrey-Smith) finds the two-fold characterization unsatisfactory and will not consider the second subdivision a legitimate one. Finally, a third view could be that of O'Malley & Powell (2016), who agree that the two-fold characterization is problematic and does not give us a coherent category. However, they argue to drop the whole idea of uniformity and a natural kind view and instead see METs as "turning points" with different ecological and evolutionary aspects that do not need to fit into a coherent category. Under this view, there is no need to talk about sub-kinds of a coherent natural kind. My aim is not to settle this debate, but rather, to add more

As Boyed (1999) has shown in the case of evolving biological species, the HPC view also helps us to avoid the rigid, essentialist, and inflexible usage of natural kind concepts because it does not consider homeostatic properties to be unique and exclusive to that natural kind. Instead, natural kind terms in the biological sciences become more flexible and without sharp boundaries. They can lose or gain some individual properties insofar as they keep a holistic cluster and play a useful explanatory role in our scientific theories. In the same way, it makes sense to accommodate for flexibility in our concept of MET and gray areas that connect them with other significant possibility-expanding evolutionary changes without dropping the concept. It is, however, important not to drop or give up on using a natural kind concept of MET and embrace a mixed and incoherent concept of it. This will amount to giving up on the explanatory aspirations that one could have about the hierarchical complexity of life. For example, information criterion cannot form a coherent category (as a natural kind or sub-kind) with fine-grained evolutionary explanatory principles. The notion of information is not well-defined here. It could involve genetic, developmental, neural-behavioral, and linguistic-cultural information. McShea and Simpson (2011) have argued that notions like complexity and the ability to survive are too vague and cannot be operationalized to form a list of major transitions with theoretical unity. The same applies to this broad notion of information. Also, Fitness values shaping the workings of these various information systems mean different things. Even if we postulate an explanatory principle for these events, there is no interesting sense in which this principle is fine-grained.

---

supporting weight to the natural kind view. My point is that if we want to keep the natural kind view and its theoretical merits, there are ways to treat major transitions as a flexible concept without sharp boundaries and acknowledge the gray areas between categories (drawing on Boyd's HPC account of natural kind). Also, I discuss the possibility of incorporating the behavioral and ecological outcomes of major transitions into its definition by the end of Sect. 4, where I discuss the dialectical method of modifying definition as the process of explanation proceeds and climbing down the ladder of explanation.

Another reason to support the restrictive and natural kind concept of METs is to avoid the possibility of conceptual confusion. This confusion could manifest itself in different ways. For example, if one uses a permissive notion of MET and still expects to draw useful explanatory generalizations among all the major events. Another confusion (and this is important for our aim) will emerge when one wants to apply transitional thinking to more specific domains of evolutionary history, like the evolution of cognition, the evolution of immune systems, or visual systems. Can one speak about major transitions in the evolution of certain traits, behaviors, and organs? In one sense, the answer is no; major transitions by definition involve the emergence of higher individuality in the biological populations, and it could not be applied to significant, possibility-expanding, and novel changes in traits or organs. In another sense, however, the answer is yes. One could talk about major transitions in the evolution of cognition and visual systems as natural-kind concepts. As discussed above, MET is a subset of novel and possibility-expanding evolutionary events that involve a defining criterion, that is, higher individuality. There is no reason why one should not discover other coherent subsets in this big set of novel and major evolutionary events that could function as natural-kind concepts. The conditions for this subset of events to reveal a natural kind are that if (1) they share homeostatic properties that could allow for a uniform definition and a list of events with theoretical unity and, (2) one could make useful explanatory generalizations from properties of one to others and consider specific and fine-grained evolutionary explanatory principles that account for them in a way that their explanation is different from explaining other random or trivial evolutionary events. In this sense, the evolution of complex biological systems and behaviors (like visual systems and learning behaviors) is explored to find qualitative leaps and possibility-expanding events that share objective similarities and could play the role of a natural kind in explanatory generalizations.[5]

The plausibility of major transitions in areas like cognition lays the ground for the next section, where we discuss different proposals in MCTs. Before that, however, we still need to discuss in more detail the second requirement that all lists or proposals of METs should meet: a fine-grained explanation of METs.

2.2. Explaining METs

After discussing the HPC view of a natural kind, it must be clear now that the first and second requirements (uniform definition and fine-grained explanation) are conceptually intertwined. If MET is not uniformly defined by a set of homeostatic properties, then it cannot delimit a unique subset of causal mechanisms (evolutionary processes) that sustain their homeostatic co-occurrence and if a unique set of fine-grained explanatory principles cannot be evoked to explain METs, then they lose their distinctive identity and remain under-explained as a category.

Okasha (2022, 3-4) has argued that the fine-grained explanations of METs should use explanatory principles and evoke evolutionary processes that are characterized in a more specific and detailed way compared to mentioning general evolutionary principles (like the fitness of cooperation and energy efficiency), which explain other normal evolutionary changes. These general principles

---

[5] It is worth noting that my claim about the MET-MCT distinction is limited. METs understood from the perspective of Darwinian individuality do not overlap with MCTs. However, if one favors a different perspective on the history of life and its transitions, they might overlap with MCTs. For example, as discussed above, the teleological transitions to a sensitive soul based on an Aristotelian view of life could overlap with MCTs (granted that the teleological transitions meet the requirements of uniform definition and fine-grained explanation). This overlap does not threaten the natural kind view. What overlaps here are the different descriptions of the same transitions. Different descriptions of transitions are desirable because different ways of looking at life and its history with different epistemic aims are desirable. Moreover, the HPC view of the natural kinds is not essentialist and is not committed to sharp boundaries. The question of what perspectives on the history of life and its major transitions are coherent, and epistemically useful, and where they overlap with each other, or other domain-specific transitions (like MCTs) lies beyond the scope of this paper.

have a role in explaining METs, but the explanations that they generate are “coarse-grained” and not “fine-grained.” Additionally, an adequate explanation of MET should involve a detailed description of transitions and stages that led to them in a way that functional and mechanical aspects of biological explanation are integrated.

One way of understanding Godfrey-Smith’s (2009, 2011) conceptualization of METs is to understand it as characterizing those specific and fine-grained aspects of evolutionary processes that generate METs. Like Michod (1999), he defends a restrictive notion of METs based on the higher-individuality criterion. He defines evolutionary individuals as members of “minimal Darwinian populations”, that is, populations that meet the minimal requirements of having heritable variation and differential fitness of their members. This makes them subjected to Darwinian natural selection. In his account, METs happen when a new Darwinian population arises, and its members are hierarchically more complex units composed of pre-existing Darwinian individuals. In his terms, MET happens when a marginal Darwinian population becomes a paradigm Darwinian population. Marginal Darwinian populations are those that approximate conditions of being minimal Darwinian populations, but they are not quite there. Paradigm Darwinian populations are those that enjoy not only minimal but also a richer set of conditions that enable them to show novel, complex, and non-trivial evolutionary changes. What are these conditions for more evolvability? Godfrey-Smith mentions conditions like how in a population reproduction works and how variations are produced, how much it is produced, and how these variations get linked to phenotype or fitness differences. For example, if the genotype in a population is robust and has a stronger force in shaping the outcome of fitness differences despite extrinsic environmental noise, then that population is closer to the paradigm Darwinian population. Or if a population shows features like the distinction between germ and somatic cells in its reproduction, then it is moving in the direction of becoming a more Darwinian population. Populations are always moving in the biological space, and they are becoming more Darwinized and de-Darwinized. Higher Darwinian individuals are members of those Darwinian populations that are built from units of a previous Darwinian population that has been marginalized or de-Darwinized (Godfrey-Smith 2011, 70-77).

Note that what we see in this account is not just a general theory about conditions of evolvability and a framework for understanding METs. This is also a fine-grained characterization of those aspects of evolutionary mechanisms or processes that explain METs. Darwinization and de-Darwinization processes in specific contexts can be evoked as fine-grained explanatory principles. Therefore, our explanation for METs is not just that natural selection selects for the fittest traits in populations, but that the Darwinizing process turned a marginal population into a paradigm one with higher individuality. The process of Darwinization presupposes the working of the process of natural selection and the survival of the fittest traits, but it is articulated at a more fine-grained level.

3- Comparing proposals in the cognitive transitions

After discussing issues around defining and explaining METs, we can now turn to the question of applying transitional thinking to cognitive evolution. There is a consensus among scholars that cognitive evolution has been marked by qualitative jumps that open new possibilities for cognitive evolvability. Different proposals about transitions in cognitive evolution have been offered (Ginsburg & Jablonka 2010, 2019, 2021, Barron et al. 2023 Hauser et al. 2002, Dennett 2017, Dehaene et al. 2022, Tomasello 2014, Graziano 2019, Dor 2023). Here, I focus on two recent, big-picture, and well-developed proposals, that is, the cognitive transition proposal offered by Barron et al. (2023) and Ginsburg and Jablonka (2019, 2021). These proposals focus on changes in the general architecture of the cognitive and nervous systems instead of domain-specific cognitive capacities like recursive syntax, use of symbols, and joint intentionality (Hauser et al. 2002, Dehaene et al. 2022, Tomasello 2014). The uniqueness of human cognition and tracing its roots in evolutionary time is not a primary concern for either. This enables these views to have a more comprehensive view of the evolution of cognition in general and to avoid anthropocentric views that tilt toward changes that make cognitive evolution appear as directed toward unique human cognition. Another reason to consider these proposals is that their authors directly refer to the MET tradition and draw links between their conceptions of transition and the ones used in the MET literature.

As mentioned above, there is one sense in which talking about major evolutionary transitions in cognition does not make sense because it does not involve higher individuality in Darwinian populations. However, there is another sense in which it makes sense to talk about major transitions in cognition, namely, if we consider MCTs a different set of possibility-expanding evolutionary events with different defining criteria and fine-grained explanatory principles compared to METs. Practically, it is in this second sense that Barron et al. (2023) and Ginsburg and Jablonka (2019, 2021) develop their proposals. As will be discussed, their proposals have a uniform definition for MCT, and there is theoretical unity in their list of major transitions. Their aim is not to use the term "major transitions" in an arbitrary way; rather, each wants to pick out events with objective similarity. However, both have their specific understanding of the relationship between their list and the MET research. Barron et al. understand Maynard Smith and Szathmàry's account of major transitions as "focused on changes in the structural basis of inheritance." Similarly, they propose to understand MCTs as structural changes in the organizational features of nervous systems or what they call the "computational architecture" of brains (Barron et al. 2023, 2). In contrast, Ginsburg and Jablonka (2021, 2019) understand Maynard Smith and Szathmàry's account to provide an account of hierarchical complexity in an "informational framework." In other words, they think that Maynard Smith and Szathmàry's account shows how the origin of new individuality is in the way information is transmitted, controlled, and interpreted. However, Ginsburg and Jablonka criticize them for having a very limited genetic-centered notion of information. Only the last transition in Maynard Smith and Szathmàry's account involves linguistic information, and the others involve genetic information. They argue that other types of information (epigenetic-cellular and neural) were ignored in their initial list or only mentioned because of the genetics that gave rise to them. So, this limited view of information missed some major transitions, like transitions in neural information (Ginsburg & Jablonka, 2019, 242). Therefore, they understand their MCT list to be a subset of (or at least overlap with) the MET list and add the missed neural transitions. Transitions in ways of transmission and control of neural information do not involve higher individuality. It seems that Ginsburg and Jablonka are comfortable adding their coherent MCT list as a subset to METs, even though that might mean that the general MET list will become miscellaneous.[6]

Now, both Barron et al. and Ginsburg and Jablonka do not discuss the problems with the two-fold characterization of MET found in Maynard Smith and Szathmàry's account. However, the former understands their proposed list of MCTs to be a different subset of possibility-expanding evolutionary events that is similar to METs, while the latter considers their list of MCTs to overlap

---

[6] Another possible (maybe more charitable) interpretation of Ginsburg and Jablonka's work could be that they accept both the informational aspect of Maynard-Smith and Szathmàry's list and types of learning as two equally legitimate perspectives for looking at the history of life and cognition and its major transitions. These two lists of transitions could be mapped onto each other (or partially overlap) instead of one being a subset of the other. Even under this interpretation, my point still holds. If one is interested in the merits of the natural kind view, then it follows that, given the vagueness of the concept of information, the informational list is not coherent, and one should treat MET and MCT as two distinct natural kinds with no overlap.

with or be a subset of MET. Given my discussion of restrictive and permissive notions of MET, the best way to understand the relationship between METs and MCTs is to consider them as two different sets of evolutionary events with different defining criteria. However, this misunderstood way of understanding the MET-MCT relationship is not detrimental to these two proposals. In practice, both proposals meet the mentioned requirements of uniform definitions and fine-grained explanations.

In my view, METs and MCTs belong to distinct natural kind terms with different defining criteria and are apt for being explained by different fine-grained explanatory principles. Given this view, items excluded from the original Maynard Smith and Szathmàry's list could be treated in the following way: Like the transition to language, they could be described as not part of the METs but MCTs. Or like RNA-DNA transition or origin of sex, they could either be a) part of another coherent natural kind that has not been discovered yet, b) they could be incorporated as the downstream evolutionary effects (e.g. behavioral, ecological outcomes) of other known METs c) they could be part of no coherent category. In the case of the last option, this does not mean that this important possibility-expanding evolutionary event is ignored. Rather, this kind of event still needs careful evolutionary considerations.

3.1. Defining MCTs

Barron et al. (2023) define MCTs as those changes in computational architecture and nervous system that open space for later cognitive evolvability. They understand computational architecture to be the specific way brains process information, and it could be mapped into the organizational features of the nervous system. Drawing on computationalism in cognitive science, Barron et al. hold that nervous systems could be understood as information processing systems whose basic aspects of operations, representations, memory, and control flow work together to build up their computational processes. Especially important for them are "memory and control flow." Control flow is that aspect of the information processing system that determines how information is processed, transformed from sensory input to motor output, and preserved to later shape behavior. It also determines how different parts of the nervous system coordinate together. Control flow is the computational component of their framework, and it is coupled with a connectomics component, that is, the mapping of connections in the nervous system. Drawing on connectomics-based analysis of the brain (Shih et al. 2017), they argue that one can map the organizational features of different brains based on their control flow. For example, Drosophila's brain represents a recurrent flow between two processing centers (motor and premotor loop) while the nematode's nervous system could best be described as a feed-forward system. They call this difference (fleshed out in terms of control flow and connectomic mapping) a difference in computational architecture and consider it to determine specific cognitive capacities that are evolvable under that architecture (Barron et al. 2023). Major transitions in cognition for them are the structural changes in the computational architecture of the brain, and they highlight five MCTs as especially important:

Barron et al.'s (2023, 4) transition proposal (5 types of computational architecture)

1. decentralized computational architecture (e.g., brain of a hydra)
2. centralized computational architecture (e.g., a flatworm's brain)
3. recurrent computational architecture (e.g., an insect's brain)
4. laminated computational architecture (e.g., a bird's brain)
5. reflective computational architecture (e.g., a human brain)

In contrast to Barron et al.'s computational architecture approach to cognition and brain, Ginsburg and Jablonka (2019, 2021) have a learning-based approach to cognition. They define cognition based on learning and provide a very broad definition of learning that can include the behavior of both neural and non-neural organisms. For them, learning is any "process leading to an experience-dependent behavioral response of a system." So, any system that can encode a stimulus, and store and process it in a way that shapes its behavior could be described as a learning system (Ginsburg & Jablonka 2021, 2). In this broad definition, one could talk about bacterial, cellular, epigenetic, and synaptic learning. Neural transitions for them are changes in learning mechanisms "that integrate, evaluate, and store" neural information and "coordinate the actions of the organism" (Ginsburg & Jablonka 2021, 3). They understand neural information in a biological framework, and it can be considered as non-genetic information that enables an organism with a nervous system to better control its body and its behavior with the environment. According to this definition of MCT, they propose five major MCTs:

Ginsburg & Jablonka proposal: (2021, 3)

1. The transition from non-neural to neural organisms (synaptic learning by habituation and sensitization)
2. The transition to animals with a central nervous system (limited associative learning or LAL)
3. The transition to animals with unlimited associated learning (UAL) with hierarchically organized brains enabling mental representations (subjectively experienced mappings of world, body, and prospective actions)
4. The transition to imaginative animals with the additional hierarchical level of neural and mental representation.
5. The transition to symbolizing, culturally learning humans

Habituation and sensitization are the ways in which an organism modifies its threshold for responding to stimuli. It can ignore irrelevant stimuli or be more sensitive to relevant ones. This type of learning occurs in single-celled organisms and is enabled by epigenetic molecular mechanisms like chromatin marking and preserved molecular records, i.e., cellular memory. In the first transition to synaptic learning, the same habituation-sensitization learning style becomes more plastic and flexible with the introduction of synaptic memory and learning. In other words, synaptic mechanisms in decentralized nervous systems enabled more rapid, targeted, specialized information transmission between stimuli and response. This enabled more complex coordination of activities and new ways of habituation and sensitization in the relation between stimuli and responses. This led to more behavioral flexibility and plasticity in learning. Likewise, in transitions to limited and then unlimited associative learning (LAL and UAL), higher plasticity and behavioral adaptability are selected. In both cases, the pairing between stimuli and response is learned without habituation-sensitization. The difference is that LAL cannot encode compound and multimodal stimuli and cannot pair stimuli and responses with a temporal gap, while UAL can form a representation of compound stimuli and has a dedicated brain area for memory and recalling these associations. UAL is an open-ended learning that enables organisms to ascribe "motivational values" to "multimodal compound stimuli and action patterns" and use them for future learning (Ginsburg & Jablonka 2021, 3-6).

In each neural transition, what is selected is the new types of learning, and the nervous system enables that type of learning, follows, and accommodates the new type of learning. Driven by selection for learning, nervous systems, as developmentally plastic systems, go through structural changes in their organizational features to enable new types of learning, and genetic accommodations fix these changes. For example, UAL as a type of learning is both a biomarker for and a driver of the evolution of its corresponding nervous system, which Ginsburg and Jablonka call the "minimally conscious" nervous system. This type of nervous system has dedicated brain areas for integrating information from different cognitive and perceptual sources, the ability to select and attend to some information and inhibit unnecessary pathways/processes, dedicated memory areas, and regions for mapping the body and the world, and a distinct sense of self and intentionality. They argue that these features have mostly evolved during the Cambrian explosion in arthropods and vertebrates and as a result of cumulative and gradual selection for UAL (Ginsburg & Jablonka, 2019, 191-240).

In Sect. 4, I will assess and evaluate their different ways of defining and explaining MCTs. However, just by going through their respective lists, it is clear that both lists are coherent, given their uniform characterization of cognition. Each transition in their list corresponds to their defining criteria, namely, computational architecture and ways of learning and processing neural information. Moreover, learning is important for both control flow and processing neural information, and the transitions in their list can roughly be mapped onto each other. For both, the first transition led to organisms with a decentralized nervous system, the second involved the emergence of a central nervous system, and the remaining three all involved more hierarchical complexity and layers of the brain. However, they differ in the terms they use for describing the transitions in cognition. Barron et al. use a hybrid computational-neurological terminology (types of computational architecture) while Ginsburg and Jablonka talk in terms of cognitive and behavioral capacities understood as different "types of learning" (e.g., habituation-sensitization, LAL, UAL). Major transitions for them are transitions in learning types. Changes in enabling nervous systems are supposed to follow and accommodate them.[7] In contrast, for Barron et al., major transitions are changes in organizational features of nervous systems, and changes in types of learning are made possible by them. In other words, MCT is seen at two different levels: as changes in biological structure or changes in behavioral outcomes of that biological structure.

3.2. Explaining MCTs

---

[7] For a detailed discussion of neural transitions and their effects, see Ginsburg and Jablonka (2019, Ch 6-8). Also, for a detailed discussion of the transition to imagination and linguistic learning see Dor and Jablonka (2010), (2014), Dor (2015), and Jablonka, Ginsburg, and Dor (2012).

As discussed above, in the case of METs, fine-grained explanations are those that highlight specific aspects of natural selection and developmental and ecological factors that drive major transitions. Similarly, in the case of MCTs, fine-grained explanations should highlight specific aspects of natural selection and other evolutionary factors that drive major transitions in cognition. The specific aspect of natural selection and evolutionary processes that is highlighted by Barron et al. is “resource constraint” and “energy efficiency.” Drawing on Calcott’s (2011) idea that similarity among properties of METs should be understood as analogous responses of organisms to environmental constraints, they argue that MCTs are driven by resource constraints as a universal environmental constraint. Since “brains are energetically expensive and have evolved under tight metabolic constraints,” one could expect selective pressure to favor those variations that enabled brains to do the same cognitive tasks more efficiently. This would have changed the computational architecture of the brains, which in turn later on made cognitive evolvability possible. In other words, increasing cognitive complexity in the evolution of brains is a byproduct of the selection for their energy efficiency. For example, the decentralized architecture allowed for more efficient and quick coordination across different body parts. Centralization allows for the efficient integration of information coming from multiple sensory organs, while recurrent architecture allows more energy efficiency than feed-forward networks. (Barron et al. 2023, 6). Barron et al. compare their “resource explanation” to similar useful strategies used for explaining the evolution of vascular plants and cardiovascular systems in animals (Niklas 2004, Monahan-Earley et al. 2013).

Ginsburg and Jablonka provide their different explanatory strategy. For them, MCTs happen because of the selective pressure of an “extended notion of selection” and fitness. They offer an extended view of selection drawing on George Price's (1971) notion of “sample selection.” Their explanatory strategy has two components. Firstly, their "evo-devo" approach highlights the role of "robustness and developmental plasticity" in evolution. Secondly, they use an extended view of selection and fitness called "sample selection." Simply put, sample selection is choosing a subset of variations based on some value criteria without the need for multiplication and replication characteristic of Darwinian populations. Darwinian selection happens in Darwinian populations. Sample selection and Darwinian selection are both included in a general concept of selection. The behavior and development of most organisms show instances of sample selection with their capacity for developmental plasticity. When faced with a challenge, organisms start random explorations and solutions and then stabilize the most efficient ones (they call it the principle of exploration-stabilization). Sample selection and this principle give organisms the ability to produce "adaptive improvisations." In the case of nervous systems, this selection plays a central role in their development and also the functioning of brains and learning (happening at multiple levels, like cellular and synaptic connections and neural selection of networks). Sample selection (with the help of developmental plasticity) drives more adaptive types of learning in neural organisms. The same sample selection explains the emergence of MCTs (Ginsburg & Jablonka 2019, 240-245). For example, the emergence of UAL is the result of cumulative sample selection, and these changes are accommodated by genetic fixation and cumulative changes in the structure of the enabling nervous system for that type of learning.

One clear similarity in the explanatory strategies of both accounts is their focus on energy efficiency. Sample selection increases the adaptiveness of organisms and leads to an extended notion of fitness that they describe based on Karl Friston's "principle of free energy." According to this principle, organisms are selected to become more efficient and minimize the free energy and disorder in the equilibrium between their body and the environment. In both accounts, energy efficiency plays a central role. According to Barron et al., natural selection selects for brains that are more energy efficient. In Ginsburg and Jablonka's view, sample selection pushes for minimizing free energy and optimizing the learning interaction between organisms and the environment.

4- Assessing the success of the two proposals

Let us start with assessing their differences in explaining MCT, and then talk about their differences in defining it. Barron et al.'s explanatory strategy and that of Ginsburg and Jablonka are successful in different ways. They are both fine-grained because they highlight different specified aspects of the evolutionary process. Ginsburg and Jablonka's account is better developed in highlighting the role of developmental plasticity and exploration-stabilization processes. These factors play a central role in the evolution of brains, given how important they are in the development and functioning of the brain (Anderson 2014). Therefore, one could praise Ginsburg and Jablonka's work for fleshing out this aspect and criticize Barron et al.'s approach for leaving this aspect of the proposal underdeveloped.[8] However, Barron et al.'s account starts at a higher level of abstraction. They abstract away from the behavioral outcomes of the different types of nervous systems. The explanandum for them is not primarily the learning and behavioral outcomes of MCTs, which "might be beneficial for downstream lineages" (Barron et al., 2). Instead, they are interested in structural changes in the organization of the nervous system. While for Ginsburg and Jablonka, major transitions are seen in terms of their behavioral and learning outcomes. In their account, changes in nervous systems are conceived to be driven by the transitions in learning. For Barron

---

[8] To be fair, Barron et al. (2023) treatment of the issue is within the word count of a journal paper. One hopes that if they have more space to develop the implications of their views (e.g. book chapters or books), they could flesh out the importance of evolutionary-developmental factors in their account more.

et al., transitions in learning could be made possible after structural changes in the nervous system are already in place by selection for energy efficiency.

It might be tempting to argue that they talk about two different sets of events. However, there is no unbridgeable gulf between changes in biological structures (nervous system) and later changes in the behavioral outcomes (types of learning) related to these biological structures. Their evolutions act back on each other. Evolved behavioral outcomes determine what structure is energy efficient, and the structure determines what level of developmental and cognitive plasticity and improvisation is allowed. For example, the emergence of centralized computational architecture was driven by selective pressure for energy efficiency and resource constraints. However, the emergence of this neurobiological structure allowed the first organisms equipped with it and their downstream lineages to evolve the LAL type of learning. This type of learning changed the evolutionary and ecological landscape and set the ground for the next possible change in structure. It makes more sense to consider their different proposals as discussing the same set of events from different perspectives. In evolution, behavioral outcomes of biological structures are as real as the structures themselves (I will come back to this point at the end of this section, where I discuss how to reconcile their different ways of defining MCTs).

In their explanatory strategy, Barron et al. focus on the immediate fitness value of computational architectures (biological structures) while Ginsburg and Jablonka highlight the fitness value of types of learning (outcomes of those structures) that might emerge later on and in the distant future. In other words, in the former, cognitive complexity is not directly selected in transitions and is a byproduct of selection for energy efficiency. In Ginsburg and Jablonka, transitions involve a direct selection for cognitive complexity, granted that one understands selection and fitness in an extended way. As Barron et al (2023, 6-7) have pointed out, variations in computational architecture were cumulatively selected because of their immediate fitness values (less energy-expensiveness). Not because of the distant value of cognitive outcomes and types of learning, that is, new and more complex cognitive capacities that evolved in the downstream of the lineages. Discounting the importance of immediate fitness values and counting too much on the fitness values of distant outcomes threatens to turn a scientific biological explanation into a prospective type of explanation (Trestman 2013). For example, in Ginsburg and Jablonka's strategy, the fitness value of “distant” functional outcomes (like UAL) is used to explain why the nervous system enabling UAL was selected in the first place. However, the question of how much Ginsburg and Jablonka’s explanatory strategy falls prey to prospective explanations is more of an empirical question than a conceptual one. It is known that in the evo-devo approach, evolution does not always have to wait for the right biological structure and its genetic composition to be in place to push for innovations. Sometimes, behavioral, developmental, and epigenetic plasticity in combination with environmental demands put a structure in place, and then genetic accommodations fixate or assimilate it in the population. Unless we know how distant in the future downstream lineages we are talking about, it is not clear whether types of learning were really in the distant future or present (in some primitive forms) and contributing to the immediate fitness value of these organisms. In other words, even though the prospective explanation is a legitimate worry raised by Barron et al. (2023), it is not clear whether Ginsburg and Jablonka are guilty of it. Their explanatory strategies do not need to replace each other. They both can contribute in different proportions, given the empirical details of each transition.[9] While selection based on resource

---

[9] Some other authors have noticed a discrepancy between the evolution of nervous systems and the evolution of types of learning in Ginsburg and Jablonka's accounts (Browning & Veit 2021). For example, the relationship between the evolution of UAL and its enabling system is not clear in their work. On the one hand, their claim is that UAL is just a positive and not a negative biomarker for the presence of its enabling nervous system. On the other hand, they treat UAL as the driving force for the emergence of its enabling system. In that case, UAL is a necessary and sufficient marker for its enabling system. One reading of their work would be to understand UAL and its enabling system as separate architectures with distinct evolutions. In that case, Ginsburg and Jablonka cannot use the selective pressure

constraints works on the immediate fitness of the organization of the nervous system, the sample selection pulls from the future based on the fitness of types of learning.

Explanation of some transitions could be exhausted by resource constraints, some by Ginsburg and Jablonka's sample selection, and some (or maybe all) by a combination of both. There might be cases in MCTs where types of learning evolved only after the computational architecture was in place. This option is emphasized in Barron et al.'s view, which holds that after each transition, the lineage might not have "gained any new cognitive functions." Instead, new computational architecture opens up a new "phenotypic space" that can be evolved into "new cognitive capacities and new forms of learning." These new adaptive capacities and behaviors "do not explain why these transitions occurred but explain in part how post-transition architectures later become entrenched" (Barron et al. 2023, 7). This insight by Barron et al. could be improved if one adds that the new adaptive capacities and forms of learning also partially determine what energy efficiency means in the evolutionary and ecological landscape.

Now, let us assess their different definitions of MCTs. Barron et al. have a hybrid computational-neural organization criterion in defining cognition and MCTs. They deliberately avoid describing nervous systems and their MCTs in terms of behavioral outcomes. Instead, they use neurobiological levels of description. In contrast, Ginsburg and Jablonka have a learning-centered definition of cognition and MCTs.

MCTs for Barron et al. are changes at the level of biological structure (nervous system), and for Ginsburg and Jablonka are changes at the level of behavioral and learning outcomes enabled by

---

for different types of learning (e.g. LAL, UAL) as the driving force behind the evolution of their enabling systems. However, as mentioned above, the question of how separate types of learning and nervous systems have evolved and how independent their respective explanations are is more of an empirical question than a conceptual one.

those structures. Again, I believe these two definitions should be considered as supplementing instead of competing. As mentioned in the Sect. 2.1, authors like Herron (2021) have argued that METs should not be defined primarily based on their outsized ecological and evolutionary outcomes. Rather, they should be characterized firstly based on changes in the biological structure, that is, the hierarchical complexity of life. I agreed with this judgment and argued that both MET and MCT should be understood in a restrictive and theoretically fruitful way. However, our definitions of MET and MCT will remain incomplete and too abstract if we stop at this stage. A better way to put this judgment is to say that MET and MCT are primarily defined in terms of the changes in biological structure (hierarchy of life or structure of nervous system) and not in terms of their ecological and behavioral outcomes. In the next step, we need to incorporate the behavioral and ecological outcomes of these transitions in their definitions. This method of self-modifying definitions, as the process of explanation progresses, is well-known in the philosophical tradition of dialectical thinking and works of philosophers like Hegel and Marx. Biologists like Richard Lewontin and Richard Levins have hailed and welcomed using dialectical thinking in biology (Lewontin & Levins 1985). We start with an explanatorily useful definition of MET or MCT at a higher level of abstraction. Later on, as we gradually climb down the ladder of abstraction, we add more features to the phenomenon and modify our definitions to approximate a more concrete picture of reality. In the dialectical method, changing our definitions and adding new elements to them as the process of explanation proceeds is not a logical inconsistency but a useful explanatory strategy. The right question to ask here is not whether Barron et al.'s definition is correct or that of Ginsburg and Jablonka. Both are correct. The right question is with which we should start at the right level of abstraction, and what should be added later. Moreover, drawing on the HPC view of major transitions, one can see that in all MCT cases, we witness structural changes in brain organization (control flow and connectome), energy efficiency, and types of learning as co-occurring properties. If these accounts are adequate, one can hope to form useful explanatory generalizations and infer the properties of other MCTs from learning more about some.

5- Conclusion:

In this paper, I defended four claims. Firstly, one explanatory fruitful way to understand METs is to understand them as natural kinds according to the HPC view of natural kinds. Secondly, in this restrictive sense, it makes sense to talk about major transitions in cognition, but METs and MCTs are two separate subsets of possibility-expanding evolutionary events with different defining criteria. Thirdly, Barron et al. (2023) and Ginsburg and Jablonka (2019, 2021) are two successful examples of using the transition-oriented approach for cognitive evolution. They both provide two different coherent lists of MCTs with different definitions of cognition and MCTs. Finally, their different definitions and explanations could be seen as supplementing, not competing.

In conclusion, I will consider one aspect that transitional thinking in both of the offered proposals can be further developed. Robin et al. (2021) have argued that METs have been defined one-dimensionally, and the ecological effects of the major transitions need to be incorporated in their definition. This is true for MET as much as for MCT. Given that evolution is a contingent multifactorial process, one has to incorporate the role of Darwinian selection in it as much as the role of developmental and ecological factors. This aspect is not as well-developed in either of the discussed proposals. Ginsburg and Jablonka (2019) discuss the role of LAL and UAL as "adaptability drivers" in the ecological context of the Cambrian explosion. However, this aspect has not systematically been integrated into their views of MCT. Especially, in the case of Barron et al.'s explanatory strategy (resource constraint explanation), what could count as energy efficiency is highly dependent on the ecological landscape of organisms and how it changes.

Moreover, in Barron et al., the role of active niche construction and ecological engagement of organisms and how it affects the evolution of their nervous system is not well-developed. Developing an evo-devo-eco account of cognitive evolution is a promising horizon that needs to be further explored.

Bibliography

Anderson, M. (2014). *After phrenology: Neural reuse and the interactive brain.* Boston: MIT Press.

Barron, A. B., Halina, M., & Klein, M. (2023). Transitions in cognitive evolution. *Proc Biol Sci. 2023 Jul 12; 290(2002)*, 20230671.

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. (Ed.), *Species: New interdisciplinary essays* (pp. 141–185). Cambridge, MA: The MIT Press.

Browning, H., & Veit, W. (2021). Evolutionary biology meets consciousness: an essay review of Simona Ginsburg and Eva Jablonka's The Evolution of the Sensitive Soul. *Biol Philos 36, 5*. doi:https://doi.org/10.1007/s10539-021-09781-7

Buss, L. (1987). *The Evolution of Individuality.* Princeton, NJ: Princeton University Press.

Calcott, B. (2011). Alternative Patterns of Explanation for Major Transitions. In B. Calcott & K. Sterelny, *The Major Transitions in Evolution Revisited* (pp. 35-52). Cambridge, MA, MIT Press. doi:, https://doi.org/10.7551/mitpress/9780262015240.003.0003,

Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: a hypothesis about human singularity. *Trends Cogn. Sci. 26*, 751–766.

Dennett, D. (1996). *Kinds of minds toward an understanding of consciousness.* New York: Basic Books.

Dennett, D. (2017). *From bacteria to Bach and back.* New York, NY: W. W. Norton & Company.

Dor, D. (2015). *The Instruction of Imagination: Language as a Social Communication Technology.* Oxford: Oxford University Press.

Dor, D. (2023). Communication for collaborative computation: Two major transitions in human evolution. *Phil. Trans. R. Soc. B 378, 20210404.* doi:doi:10.1098/rstb.2021.0404

Dor, D., & Jablonka, E. (2010). Canalization and plasticity in the evolution of linguistic communication. In R. K. Larson, V. Deprez, & H. Yamakido, *The Evolution of Human Language,* (pp. 135–147). Cambridge: Cambridge University Press.

Dor, D., & Jablonka, E. (2014). Why we need to move from gene-culture co-evolution to culturally-driven co-evolution In D. Dor, C. Knight, & Lewis, *Social Origins of Language* (pp. 15–30). Oxford: Oxford University Press.

Ginsburg, S., & Jablonka, E. (2010). The evolution of associative learning: a factor in the Cambrian explosion. *J. Theor. Biol. 266*, 11–20.

Ginsburg, S., & Jablonka, E. (2019). *The evolution of the sensitive soul: Learning and the origins of consciousness. (A. Zeligowski, Illustrator).* Cambridge, Massachusetts: The MIT Press. doi:https://doi.org/10.7551/mitpress/11006.001.0001

Ginsburg, S., & Jablonka, E. (2021). Evolutionary transitions in learning and cognition. *Phil. Trans. R. Soc. B 376*, 20190766.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection.* Oxford: Oxford University Press.

Godfrey-Smith, P. (2011). Darwinian Populations and Transitions in Individuality. In B. Calcott & K. Sterelny, *The Major Transitions in Evolution Revisited* (pp. 65–82). Cambridge, MA: The MIT Press.

Graziano, M. (2019). *Rethinking consciousness: a scientific theory of subjective experience.* New York, NY: W. W. Norton & Company.

Hauser, M., Chomsky, N., & Fitch, W. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science 298*, 1569–1579.

Herron, M. D. (2021). What are the major transitions? *Biol. Phil. 36, 2.*

Jablonka, E., Ginsburg, S., & Dor, D. (2012). The co-evolution of language and emotions. *Philosophical Transactions of the Royal Society B 367*, 2152–2159. doi:doi:10.1098/rstb.2012.0117

Lewontin, R., & Levins, R. (1985). *The dialectical biologist.* Cambridge: Harvard University Press.

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance.* Cambridge, MA: Harvard University Press.

McShea, D. W., & Simpson, C. (2011). The Miscellaneous Transitions in Evolution. In B. Calcott & K. Sterelny, *The Major Transitions in Evolution Revisited* (pp. 19-35). Cambridge, MA: MIT Press.

Michod, R. (2007). Evolution of individuality during the transition from unicellular to multicellular life. *Proc Natl Acad Sci 104*, 8613–8618. doi:https://doi.org/10.1073/pnas.07014 89104

Michod, R. (2011). Evolutionary transitions in individuality: multicellularity and sex. In B. Calcott & K. Sterelny, *The Major Transitions in Evolution Revisited* (pp. 167–197). Cambridge: MIT Press.

Michod, R. E. (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality.* Princeton, NJ: Princeton University Press.

Michod, R., & Herron, M. (2006). Cooperation and conflict during evolutionary transitions in individuality. *J Evol Biol 19:*, 1406–1409. Doi:https://doi.org/10.1111/j.1420-9101.2006.01142.

Monahan-Earley, R., Dvorak, A., & Aird, W. (2013). Evolutionary origins of the blood vascular system and endothelium. *J. Thromb. Haemost. 11*, 46–66. doi:doi:10.1111/jth.12253

Niklas, K. (2004). Computer models of early land plant evolution. *Annu. Rev.. Earth Planet. Sci. 32*, 47–66. doi:doi: 10.1146/annurev.earth.32.092203. 122440

O'Malley, M., & Powell, R. (2016). Major problems in evolutionary transitions: how a metabolic perspective can enrich our understanding of macroevolution. *Biol Philos 31*, 159–189. doi: https ://doi.org/10.1007/ s1053 9-015-9513-z

Okasha, S. (2022). The Major Transitions in Evolution - a philosophy-of-science perspective. *Frontiers in Ecology and Evolution, 10, Article 793824*. doi:https://doi.org/10.3389/fevo.2022.793824

Price, G. R. (1971/1995). The nature of selection. *Journal of Theoretical Biology 175*, 389–396. doi:10.1006/jtbi.1995.0149

Queller, D. (1997). Cooperators since life began. *Q Rev Biol 72:* 184–188. doi:https://doi.org/10.1086/41976 6

Robin, A., Denton, K., Horna Lowell, E., Dulay, T., Ebrahimi, S., Johnson, G., et.al. (2021). Major evolutionary transitions and the roles of facilitation and information in ecosystem transformations. *Front. Ecol. Evol. 9:711556*. doi:doi: 10.3389/fevo

Shih, C. T., Sporns, O., Yuan, S., Su, T., Lin, Y., Chuang, C., . . . Chiang, A. (2017). Connectomics-based analysis of information flow in the Drosophila brain. *Curr. Biol. 25*, 1249–1258. doi:doi:10.1016/j.cub. 2015.03.021

Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution.* Oxford, UK: Oxford University Press.

Solari, C. A., Nedelcu, A. M., & Michod, R. E. (2003). Fitness and complexity in volvocalean green algae. In H. Lipson, E. K. Antonsson, & J. R. Koza, *Computational Synthesis: From Basic Building Blocks to High-Level Functionality* (pp. 218–225). Menlo Park, CA: AAAI Press.

Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proc Natl Acad Sci.* doi:doi.org/10.1073/pnas.14213

Tomasello, M. (2014). *A natural history of human thinking.* Cambridge, MA: Harvard University Press.

Trestman, M. (2013). The Cambrian explosion and the origins of embodied cognition. *Biol. Theor. 8*, 80–92. doi:10.1007/s13752-013-0102-6