teorema
Vol. XLIII/3, 2024, pp. 39-62
ISNN 0210-1602
[BIBLID 0210-1602 (2024) 43:3; pp. 39-62

Predictive Architectures Cannot Be Modular

Sergio Cermeño-Aínsa

ABSTRACT

Drayson (2017) explores the relationship between predictive and modular architectures of the mind and concludes that predictive architectures must exhibit some kind of modularity. To do so, Drayson discusses two requirements of predictive architectures that seem to conflict with modular architectures: the continuity claim, the idea that cognition and perception rest on a continuum, and the non-isolation claim, the idea that no brain processes are informationally isolated. Although these features seem to repel modular architectures, Drayson finds reasons for reconciliation. In this paper, I explain such reasons and provide difficulties in Drayson's argumentation. I conclude that there is no place for reconciliations.

KEYWORDS: *Predictive Architectures; Modular Architectures; Continuity Claim; Non-isolation Claim; Cognitive Penetrability; Markov Blankets.*

RESUMEN

Drayson (2017) argumenta que las arquitecturas predictivas deben exhibir algún tipo de modularidad. Para ello, analiza dos requisitos de las arquitecturas predictivas: la afirmación de continuidad, la idea de que la cognición y la percepción se ubican en un continuo, y la afirmación de no aislamiento, la idea de que ningún proceso cerebral está informativamente aislado. Aunque estas características parecen repeler las arquitecturas modulares, Drayson encuentra razones para la reconciliación. En este artículo, explico dichas razones y proporciono serias dificultades a su argumentación. Concluyo que no hay lugar para reconciliaciones.

PALABRAS CLAVE*: arquitectura predictiva; arquitectura modular; afirmación de continuidad; afirmación de no aislamiento; penetrabilidad cognitiva; manta Markov.*

The type of architecture that rules mental processes is under discussion. Traditionally, modular architectures have posited compact, mandatory, functionally isolated and unidirectional (bottom-up) processing mainly for perceptual (and linguistic) systems. This functional independence suggests that perceptual systems passively receive and process external information alien to the operations performed by other systems [Fodor (1983); Pylyshyn (1999); Raftopoulos (2009); Firestone and Scholl (2016)]. In contrast, predictive architectures advocate flexible, plastic, functionally integrated and bi-directional (bottom-up and top-down) processing for all mental systems. Predictive architectures represent a significant departure from traditional thinking, viewing perception not as a passive receptor of information but as an active inferential process. The idea is that higher-level cognitive systems actively anticipate the information processed by lower levels. This is achieved through some form of Bayesian inference — the higher brain generates hypotheses about how the world is, and these hypotheses are adjusted based on incoming information. Predictive feedback is then compared with the incoming information to refine the feedback signals and generate the most suitable representation of the incoming signal [Hohwy (2013); Friston (2010); Clark (2013)]. In recent years, predictive architectures have been gaining popularity, thus challenging the predominance of modular architectures. Many believe that a paradigm shift is underway.

Although there seem to be no reasons for reconciling these different accounts, Drayson (2017) argues that the tension could be relieved. This tension arises for two reasons. Firstly, predictive architectures assume that perception and cognition are part of a continuum implemented by the same type of mechanisms, challenging the idea proposed by modular architectures that there is a clear boundary between perception and cognition (the continuity claim). Secondly, according to predictive architectures, no process in the brain operates in isolation; thus, the modular assumption that some parts of perceptual processing are functionally isolated from cognitive processing is no longer tenable (the non-isolation claim). Drayson (2017) questions these two claims and argues that predictive architectures can coexist with some form of modularity. This paper rejects this view. I argue that predictive and modular architectures are antagonists — there cannot be modules in predictive architectures. The paper is structured as follows: The first section clarifies the notion of modularity that Drayson refers to. The subsequent section provides a brief explanation of predictive architectures. Next, I explain Drayson's conciliatory proposal by highlighting the two claims (continuity and non-isolation) widely adopted by predictive architectures. Then, I present some challenges to Drayson's argument and conclude that there is no room for reconciliation — if minds are predictive, they cannot be modular.

I. DIFFERENT NOTIONS OF MODULE

In "Modularity of Mind", Fodor (1983) introduced an influential view on human mind functioning. Fodor suggests that low-level peripheral systems, such as perceptual and linguistic systems, operate as modular entities, while high-level perception and cognitive systems do not. Fodor outlines up to nine properties of modular systems, but in later writings [Fodor (2000)], he emphasises informational encapsulation as the fundamental property of modularity. A system is informationally encapsulated when it operates independently, without relying on information processed by other systems, particularly higher-level cognitive systems.

Since their initial formulation, modular architectures have undergone various transformations. For example, proponents of massive modularity [Carruthers (2006); Sperber (2001); Barrett and Kurzban (2006)] argue that the entire human mind, including high-level cognitive systems, is modular. However, the thesis of massive modularity requires a notion of module different from that proposed by Fodor. Despite significant explanatory costs, proponents of massive modularity (mostly evolutionary psychologists) are forced to reject or at least weaken informational encapsulation as a robust criterion for modularity [Coltheart (1999); Carruthers (2006)].[1]

Beyond the differences between these versions (essentially, but not only, in the scope of the notion of module), there is an essential common element — both coincide that a system is modular to the extent that its processes are functionally specialized and, therefore, relatively isolated from the processes carried out by other systems. This relaxed notion of module is, following Burnston and Cohen (2015), p. 132, the one embraced by Drayson [see also Beni (2022)]. This characterization allows for some overlap between modules — if necessary, some information can permeate between them —, and softens the informationally encapsulated criterion at the expense of weakening the explanatory force of the notion of module [see Stokes and Bergeron (2015)]. Indeed, the more unencapsulated a module is, the more it will need to compute the information processed by other modules, thus losing the explanatory power provided by modular impermeability[2]. However, even this more flexible characterization conflicts with predictive architectures — modules cannot accept that perception and cognition lie inseparably on a continuum (continuity claim) nor that any part of perceptual processing is functionally isolated from high-level cognitive processing (non-isolation claim). This is precisely what predictive architectures suggest.

II. PREDICTIVE BRAINS

Predictive processing (PP) reverses traditional thinking by suggesting that biological brains are not just passive receivers of sensory input but are constantly active and engaged in predicting sensory stimulation. According to predictive architectures, the brain continually generates top-down predictive models or representations of the world that constitute predictions of how the world is [Rao and Ballard (1999); Lee and Mumford (2003); Hohwy (2013); Clark (2013); Friston (2010)]. These top-down predictions facilitate and accelerate perceptual processing by reducing the need to reconstruct the environment through exhaustive bottom-up analysis of incoming sensory information [Panichello et al. (2013), p. 4]. In other words, perceptual systems rely on prior probable representations stored in memory to adjust incoming sensory inputs, so perceiving the world consists of continuously elaborating estimations about how the world is.

The perceptual brain is therefore continuously active, with cognitive and sensory systems working jointly, interactively and concurrently to predict incoming sensory stimuli. Processes such as image segmentation, surface inference, figure-ground segregation, contour integration or object recognition do not progress in a purely bottom-up serial fashion but in continuous feedforward and feedback loops that simultaneously activate the entire hierarchical circuitry of the visual system. From this perspective, the brain continuously generates expectations and predictions about the immediate future. In simple terms, brains are essentially prediction machines [Clark (2013)].

Importantly, PP is not a unified framework, as there is significant disagreement regarding the scope of its explanatory powers [Sims (2016), (2017)]. The range of explanations varies from some perceptual processes to some neurocognitive functions, all the way to encompassing all neurocognitive functions and even all biological self-organization [Sims (2017), pp. 4-5]. Despite this, predictive architectures coincide in two postulates. First, cognition and perception lie on a continuum (the continuity claim), and second, no process in the brain is functionally isolated (the non-isolation claim). These postulates seem incompatible with the existence of modules in the brain; however, Drayson claims to have found a way for reconciliation. The following section examines Drayson's compatibility arguments.

III. DRAYSON'S ARGUMENTATION FOR COMPATIBILITY

For Drayson (2017), the two crucial assumptions of predictive architectures — the lack of boundaries between perception and cognition (the continuity claim) and the lack of informational and functional isolation of perceptual processing (the non-isolation claim) — are compatible with modular architectures. Note that the two claims are closely related; if there are no boundaries between perception and cognition, then no part of perceptual processing is informationally or functionally isolated from higher-level cognitive processing and vice versa. Drayson, however, discusses the two claims separately.

III. 1 *The Continuity Claim*

When discussing the continuity claim — the idea that there is no sharp boundary between perception and cognition — Drayson provides two arguments for reconciliation. The first is that even though continuity involves no differences in Bayesian computational processes across the processing hierarchy (i.e., seeing an apple and thinking about the future involve similar computational mechanisms), the brain can still be considered a hypothesis tester and modular at the same time. Drayson appeals to the level of description [p. 7], stating that a system can be considered modular or not, depending on the level of the grain employed to describe it. At the fine-grained level of description, seeing an apple and thinking about the future are executed by the same kinds of computational mechanisms, whereas at a coarse-grained level of description, seeing an apple and thinking about the future are not continuous. Thus, perception and cognition can be continuous in fine-grained details and discontinuous in coarser ones — it all depends on the level of description. Fundamentally, Drayson is referring, following Davies (1989), to coarse and fine-grained levels of description as functionally characterized. She argues as follows:

> modularity is a matter of grain: a computational system can be modular when viewed at one level of abstraction but not when viewed at another: continuity in the fine-grained details of the information-processing is compatible with discontinuity at a coarser-grained perspective" [Drayson (2017), p. 7].

Thus, predictive and modular architectures are not mutually exclusive since similar fine-grained mechanisms may originate distinct functional modules in a coarse-grained perspective.

In the second argument Drayson argues that although the continuity claim implies that perception and cognition are at both extremes of a spectrum, and most everyday processes fall in the middle (without clear boundaries of what counts as perception or cognition), there will always be perceptual but not cognitive cases at the lower end of the hierarchy and cognitive but not perceptual at the higher end (p. 7). This seems at odds with the idea that there are ultimately no fundamental differences between perception and cognition.

III.2 *The Non-Isolation Claim*

Roughly, the non-isolation claim suggests that there is no part of perceptual processing informationally and functionally isolated from cognitive processing. This is because priors affect each level in the predictive Bayesian hierarchy from the level above. Since there is no level in the hierarchy unaffected by the preceding level, predictive architectures must entail cognitive penetration of perception. Drayson presents two arguments to refute this conclusion. The first questions the type of cognitive penetration implied by predictive architectures. Supporters of PP need to clarify the nature of cognitive states involved in cognitive penetration, whether doxastic (conscious beliefs) or subdoxastic (unconscious information represented in the cognitive system) [see also MacPherson (2017)]. If cognitive penetration requires doxastic states, PP supporters have two options: either include doxastic states as the cognitive states involved in predictions or question the real existence of such doxastic states, adopting an eliminativist approach [Dewhurst (2017)]. According to Drayson, these options do not leave much room for cognitive penetration. Leaving aside the second option, Drayson seems reluctant to think that the type of top-down predictions suggested by PP involve doxastic states.

The second argument relies on the logical property of transitivity. Predictive networks assume that each level is causally influenced by the level above: n+1 causally influences n, which in turn causally influences n-1, and so on. This, in principle, means that n+1 also influences n-1. However, Drayson argues this may not be the case, since the causal influence might only occur between correlative levels. Furthermore, the further apart the levels are in the hierarchy, the less causal influence there will be between distant levels, implying some level of isolation for perceptual processes.

Next, I raise significant objections to these arguments. I argue that if the picture drawn by predictive architectures is correct, then the idea of a modular mind should be abandoned.

IV. CONTRA-ARGUMENTATION: INCOMPATIBLE ARCHITECTURES

At first glance, Drayson makes room for two seemingly antagonistic approaches. However, there are additional reasons beyond those mentioned by Drayson that argue against reconciliation. For instance, if every part of the visual brain is subject to be modified by predictions, the idea of specialized functional modules becomes, at best, inconsistent. Furthermore, it is hard to reconcile the dynamic, active and plastic theory suggested by PP with the static, passive and constrained picture proposed by modularity. Additionally, while predictive architectures are probabilistic-based and in lineage with connectionist models [Clark (2016); McClelland (2013)] modular architectures are non-probabilistic and reject connectionism [Fodor and Pylyshyn (1988); Bechtel and Abrahamsen (1991)][3]. These reasons hinder reconciliation, but even setting aside all this, Drayson's arguments fail[4].

IV.1 *Continuity at any Explanatory Level*

For the continuity claim, Drayson first holds that taking the fine-grained details of the information processing, the continuity claim makes sense since perception and cognition are ruled by similar mechanisms. However, from a coarse-grained perspective, perception and cognition are easily differentiable and discontinuous.

However, when PP supporters describe the computational mechanisms that rule perception, action and cognition, they explicitly adopt both the fine-grained and the coarse-grained perspectives. When they claim that the brain operates according to Bayesian rules, they are explicitly connecting the underlying mechanism to its function; i.e., the function the system computes depends on how the predictive mechanism is neurologically implemented in the brain. Therefore, the variability of the fine-grained mechanisms suggested by predictive architectures, *the delicate dance between top-down and bottom-up processing* [Clark (2013), p. 9], has a counterpart in the functional coarse-grained explanation in the form of a greater or lesser contribution of cognition in the attainment of a perceptual experience. All this means that to the extent that perception is continuously modulated by cognition, perception and cognition fall, at any explanatory level, on the same continuous, and hence the boundary between perception and cognition becomes highly imprecise. The central idea of PP that perception is, in an optimal combination, continuously framed by subjects' cognitive background suggests that perception is continuous with cognition at any explanatory level — thinking about future or past events, contemplating an artwork or seeing an apple only differ in the grades of intervention of top-down cognitive processes and bottom-up driving signals (arriving from sensory channels) to obtain one or another perceptual experience[5].

Furthermore, PP holds that top-down predictions are ubiquitous and do most of the perceptual work. But, if top-down effects are persistent in the brain, and endogenous mechanisms such as attention, memory, beliefs or mood, are vital to predict the incoming input and efficiently represent environmental information, then the presence of specialized functional modules that only process incoming information makes no sense *at any level of explanation*. This being the case, we should consider our perceptual experience as modulated by cognition and inherently joined to it. The coarse-grained level of explanation in predictive architectures indicates that perception alone cannot explain sensory input and, therefore, there should be some kind of continuity between perception and cognition at any explanatory level. Predictive architectures assert that the functional specialization posited by weak modular approaches is not an intrinsic property of any particular brain region but depends on bottom-up and top-down connections among different and varied brain areas [Friston and Price (2001), p. 275]. In short, if discontinuity in modular accounts is established by postulating an exclusively bottom-up phase, and in predictive accounts no part of the brain is free from top-down generative processes, then, regarding the criteria of continuity, modular and predictive architectures are incompatible.

Finally, if modularity depends on the grain of description, and modules only make sense at the coarse-grained level, then the mind will be modular only from our functional understanding of it. At this point, modules become useful only as functional talking about the mind. In contrast, PP shows not only that seeing an apple or thinking about the future can be explained by appealing to similar fine-grained mechanisms, but also that the kind of phenomenal experience is characterised by the degree of predictive processing, the subsequent degree of prediction error (marked by the degree of environmental uncertainty) and ultimately by the extent in which perception, action and cognition are mutually intertwined. Thus, while modules would, at most, make sense only at the coarser grain of description, where perception and cognition may be seen as functionally distinct, predictive architectures make sense in both the fine and coarse-grained levels of description without the need to appeal to modules. Indeed, besides the underlying computational mechanism, PP also describes the mechanism by which perceptual inputs are top-down predicted by prior knowledge stored in memory. Thus, the coarse-grained level appears in PP, rooted in the fine-grained level. Finally, if we want a cognitive architecture drawn on reliable data about structural, anatomical, neurophysiological and functional brain features, predictive architectures provide all these data without resorting to modules[6].

The second argument considers both ends of the spectrum, where perception and cognition appear disjointed — insofar as there must be pure perceptual cases at the lower end of the hierarchy, a boundary between perception and cognition stands. One way to avoid this conclusion is by considering a continuum where perception is so tainted by cognition that pure perception (the lower end of the spectrum) has very little (if any) relevance in our everyday lives. Indeed, one can consider a continuum where perception cannot be explained without the minimal intervention of cognitive processes, a continuum where the lower end of the spectrum becomes, at least in real-world scenarios, an unlikely situation. Ultimately, if perceptual systems aim to allow organisms to respond adaptively to ecologically relevant stimuli, then the mediation of background stored knowledge becomes indispensable. This is partly because stimuli are not represented in a contextual vacuum, but are inextricably bound to internal (mental) and contextual (environmental) information. But this is also because of the informative vagueness of sensory inputs; the ambiguity of incoming information forces us to rely on prior knowledge accumulated throughout the evolutionary and developmental process to convert sensory energy into valuable information that guides behaviour [Lupyan (2015)].

One can consider non-natural cases such as the Müller-Lyer illusion (see Figure 1a), where one cannot stop seeing the lines as different even though knowing they are equal; i.e., the illusion persists even knowing that it is an illusion. This can turn it into a paradigmatic case of pure perception (a prototypical case of the lower end of the spectrum) since prior knowledge does not interfere with the visual experience. These cases have been broadly used to argue for the modular nature of perceptual systems [Fodor (1983); Pylyshyn (1999)], but appealing to the Muller-Lyer illusion to show this is not very persuasive. Simply, from the fact that previous knowledge cannot dominate low-level processing in these circumstances, it does not follow that previous knowledge cannot influence or dominate low-level processing in other circumstances — cognitive influence might not be so determinant in these cases [Prinz (2006); Ogilvie and Carruthers (2015)]. Furthermore, this can also be seen as a prominent case of cognitive penetration of low-level sensory processing. Perhaps the visual system is conditioned by the profound previous belief that the disposition of the wings indicates cues about depth and three-dimensionality (see Figure 1b)[7]. The visual system is, in this case, more prone to obey the deeply ingrained information of depth than the previous and less relevant information about the length of the lines [McCauley and Henrich (2006)]. Therefore, depending on how high-level the involved previous information is, this can be considered a case of top-down modulation, a particular case of weak cognitive penetration even though the same top-down modulation recurrently facilitates true perception [Hohwy (2017), p. 78]. All this is consistent with the PP explanation of this and other persistent illusions. Lupyan (2015), for example, argues that

> [I]nsofar as the Müller-Lyer illusion arises from the visual system attempting to represent likely real-world sources, it would be maladaptive to undo one 'illusion' while breaking the rest of vision in the process. A bit of additional evidence in the form of training allows the system to reach a globally optimal state, making accurate local predictions while maintaining globally optimal performance [Lupyan (2015), p. 558].

Thus, the illusion arises because the ambiguity of some artificial scenarios muddles visual systems forcing them to choose between hypotheses. In this case, the ambiguity is settled too early on, thus preventing the confirmation of high-level previous beliefs. This is not to say that high-level processing is absent, but that such processing is not sufficiently tested, thus generating erroneous predictions. But again, only an extra bit of information (e.g., smaller wings) or little visual training may be enough for the system to generate true perception [Rudel and Teuber (1963)].[8]

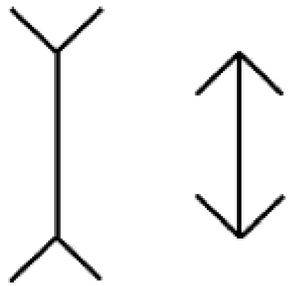

FIGURE 1A: Muller-Lyer illusion

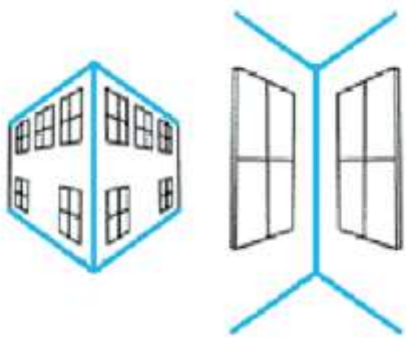

FIGURE 1B: Muller-Lyer illusion with extra three-dimensional information.

Let me consider another possible objection. One can think of a newborn whose poor or null stored knowledge of the world prevents cognition from influencing perception, e.g., via categorization. If newborns perceive but do not categorize the world, then perception and cognition are (at least during our first steps) disconnected. Infants are, therefore, positioned at the lower end of the spectrum. But do babies have genuine perceptual experiences? Arguably, newborns initially perceive a disorganized and meaningless amalgam of images, sounds or smells that over time become organized and gain meaning. Thus, if the low end of the spectrum (the supposedly pure perceptual stage) consists of this uninformative combination of sensory inputs, then that pristine perceptual experience is closer to being a mere raw sensation than genuine perception. Indeed, studies show that infants can recognize voices, faces, or odours extremely early; raw sensations begin, therefore, also very early, to be guided and modelled by memories and learned prior knowledge. The immense assortment of natural stimuli has compelled us to develop a plastic rather than rigid mental processing; in the natural world, stimuli are not simple lines of different lengths or idealized figures but rather a complex combination of different lines, colours, textures, meanings, categories and so on, whose processing requires a very diverse range of cognitive processes (sensory, emotional, cognitive and contextual information). So, if modular architecture is bounded to cases of the lower end of the hierarchy, then it ends up being informationally sterile; modules will not be very informative after all.

IV.2 *There Are No Islands in the Brain: Transitivity and Markov Blankets*

The second attempt at reconciliation suggested by Drayson is to cast doubt on the non-isolation claim — the idea that perceptual processing is not informationally and functionally isolated from high-level cognitive processing. The non-isolation claim is closely linked to the cognitive penetrability of perception. Proponents of the cognitive penetrability of perception state that the contents of perceptual experiences are influenced by high-level mental states (prominently, background knowledge, beliefs and memories). Now, since PP postulates that the brain generates top-down predictions to facilitate, guide and constrain the processing of incoming sensory input, the cognitive penetrability of perception should, according to PP postulates, be the norm. According to Drayson, however, it is far from clear that the kind of cognitive influence postulated by PP entails cognitive penetration in any relevant way. Questions such as whether cognitive influence should be exerted on late or early perceptual states, whether cognitive states should be beliefs or desires (doxastic states) or may also be just cognitive representations (sub-doxastic states), or whether the type of relationship that exerts the influence must be direct or indirect, have long been debated by philosophers interested in these topics.

The question is: what kind of cognitive states are the states involved in predictive architectures? One can consider a strong and a weak sense of cognitive states. The former refers to doxastic states — propositional attitudes like beliefs or desires accessible to consciousness and inferentially integrated. The latter refers to non-doxastic states — moods, emotions, types of personality, cognitive styles, education, learning or expectations, which are non-accessible to subjects' consciousness. In my view, PP accommodates both notions. For example, it has been reported that desiring an object makes it look closer [Balcetis & Dunning (2010)] or that believing that someone is upset with you can make you to perceive that person's facial features as particularly irritated [Siegel (2012)]. These are cases where conscious beliefs and desires (doxastic states) modulate perception. Therefore, even considering that genuine cognition is exhausted by doxastic states, there are compelling cases where doxastic states contribute to the constitution of perceptual experiences. However, perhaps a cognitive state need not be accessible to consciousness and inferentially integrated to be considered a genuine cognitive state. In most situations, we perceptually reconstruct the world without needing high-level beliefs (in the doxastic sense). The reported cases where moods, emotions, types of personality, cognitive styles, education, learning, or expectations influence perception are countless. For example, the expectation of motion alters motion perception [Sterzer et al. (2008)], negative thoughts make the world look darker [Banerjee et al. (2012)], or types of personality, cognitive styles and moods can make us perceive stimuli more vividly, oriented in different ways or even sized, shaped and coloured differently in different circumstances or different situations [Harber et al. (2011); Stefanucci and Geuss (2009); Levin and Banaji (2006)][9]. Finally, if the core of PP is the minimization of the overall prediction error, and this is achieved by a mechanism that collects the information processed from the different sensory modalities, prior experiences, expectations, stored knowledge or beliefs, then this information will be necessary to guide processing at the lower levels [Lupyan (2015), p. 547]. Thus, the doubts raised by Drayson about the nature of the cognitive part of the equation are unfounded since much of the evidence shows that the two senses of cognition (doxastic and non-doxastic) can influence perception. At this point, the relevance of the very notion of cognitive penetration could be called into question, since if cognition always penetrates perception, and there is no point where cognition ends and perception begins, then what is the point of questioning cognitive penetration? Recently, Block (2023) has argued that even if there is cognitive penetration, it is possible to defend a joint between perception and cognition, although such a joint is not marked in architectural terms, but in representational ones [for contrary positions see Quilty-Dunn (2020) and Green (2023)]. But, even though it is omnipresent, it is also possible to talk about cognitive penetration in terms of degrees to which cognition affects perception. From the point of view of PP, the degree to which high-level predictions are used to adjust lower-level representations [Lupyan (2015)]. Be that as it may, there is little doubt about the existence of cognitive penetration, whether it is relevant or not will ultimately depend on the epistemic consequences it has on the agent.

But the bulk of Drayson's challenge to the non-isolation claim goes in a different direction. Recall that predictive architectures are construed on the basis that each level of the hierarchy is causally determined by the upper level — no level is immune to receiving influences from the levels above. This suggests that higher cognitive levels should somehow shape lower perceptual ones, and therefore, our perception of the world is determined by our conception of the world. Drayson tries to avoid this conclusion by appealing to possible failures in the logical property of transitivity. Indeed, the non-isolation claim stands because predictive architectures assume that there must be causal influences across different hierarchical levels: level A+1 influences level A, level A influences level A-1, and then level A+1 influences level A-1. However, this is the case if we assume the logical property of transitivity, and Drayson suggests that this might not be the case — perhaps A+1 causally influences A, but not A-1. If probabilistic causation does not meet the transitivity requirements, then predictive architectures may not preserve the causal influences over long causal chains. Drayson reasons:

> In this way, we can accept that each level in the predictive hierarchy is causally influenced by (i.e., gets its priors from) the level above, without having to accept that each level in the hierarchy causally influences all the levels below it, or that each level is causally influenced by all the levels above it. And so it remains plausible that there are perceptual processes (lower-level processes involved in spatiotemporally precise predictions) which are isolated from cognitive processes (higher-level processes involved in abstract predictions) in the sense that the former are not causally influenced by the latter [Drayson (2017), p. 9].

First of all, it should be noted that this reasoning not only threatens the non-isolability claim but also the entire PP framework. PP considers that generative weights connecting two layers generate an activation in the layers below, which are subsequently registered to cause a prescribed pattern of activation until the activation reaches the layer just above the lowest level of input data. If, as Drayson suggests, there is no informational transitivity between more than two levels, then the causal influence exerted by the information of the upper layers should be blind to the information of the layers below, and therefore perceptual inputs could not be predicted. Draysons argument rests on an analogy with Spohn's (2009) probabilistic meteorological models: just as the weather at the turn of the last century does not make a probabilistic difference for today's weather, causal influences in PP models may also not be preserved over long causal chains. The result is transitivity failures in both cases. There are, I argue, several ways to avoid this conclusion.

First, in probabilistic causality, the influence between distant nodes may become weaker but not necessarily inexistent; some minimal influence may be preserved. For example, by adding further conditions to causal chains, the transitivity of probabilistic causality [Eells and Sober (1983)] and the transitivity of probabilistic support [Shogenji (2003); Roche (2012)] can be conserved[10]. In our context, these conditions might be wielded by the influence of lateral connections, perhaps via recurrent lateral inhibition. Indeed, although PP researchers have emphasized the importance of feedback connections, their algorithms usually capture processes implemented by lateral connections [Rao and Ballard (1999), pp. 84-86], and perhaps such connections may form a triad and establish transitive relationships [Snijders (2008)].

Second, Drayson's reasoning seems to assume that instead of a dynamical cyclic network, the predictive brain delivers a structural acyclic network (for differences, see Figure 2). Friston (2011), p. 25, for example, considers these two types of probabilistic generative models, the structural (DAG) and dynamic (DCM) causal models, and concludes that dynamic causal models better capture the essence of PP [see also Hipolito and Kirchhoff, (2019); though see Beni (2022)[11]]. Therefore, in terms of DAG, Drayson's intransitivity argument makes sense since PP would appear as a linear feedback processing, but in terms of DCM, the argument collapses. In the case of DAG, the influence between nodes follows a linear ordering, thus causing the loss of transitivity: n+1 influences n which, in turn, influences n-1. But in the case of DCM, the influence between nodes goes beyond linearity: n+1 influences and is influenced by n, which in turn influences and is influenced by n-1, which in turn influences and is influenced by n+1. All nodes are reciprocally and recurrently related. But why do the dynamics described in DCM fit better with the postulates of PP? There are several reasons. First, the dynamics represented in DCM better explain how the system goes one step ahead; and how current states predict and influence future ones (which is the essence of PP). Second, this dynamic allows the information to be recorded in memory for future deployment in case of necessity, thus providing the system with a specific autoregulatory mechanism and making it possible to preserve the information over long causal chains. And finally, DCM dynamics provide the system with a continuous rather than a discrete representation of the world. Indeed, though saccadic eye movements, for example, constitute a series of discrete fixations intercalated with rapid movements, our experience of the visual world is temporally and spatially continuous. This is because DCM's dynamics allow the system to constantly produce hypotheses about a continuous rather than a discrete world. Seen in this way, the idea of predictive brains becomes much more robust, accurate and natural without appealing to compartmentalized modules.

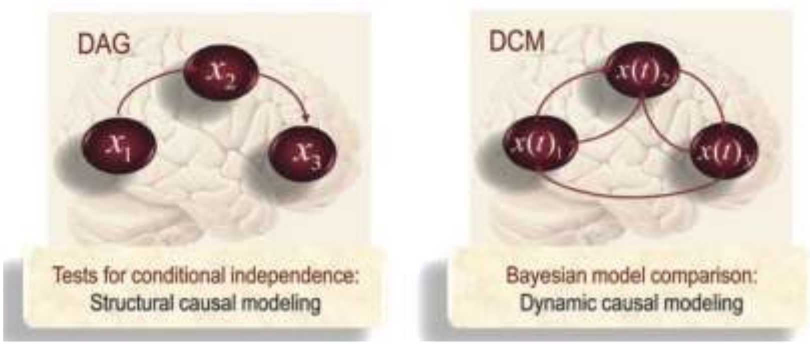

FIGURE 3: Differences between structural and dynamic causal modelling [Friston (2011), p. 25].

Finally, rather than functionally specialized modules, predictive models suggest that *functional segregation is only meaningful in the context of functional integration and vice versa* [Friston (2011), p. 15], or as recently put rather than modularly, predictive models are best understood as performing factorized mean-field approximations, a mean-field approximation being a description of the message passed from one node to another [see Parr et al. (2020)]. Let me introduce the notion of Markov blanket in the context of predictive architectures [Friston (2013)]. A Markov blanket for a node in a causal net is the node's parents, children and parents of its children (see Figure 3). Each node contains its own Markov blanket, and it is in some sense closed inside it (x's nodes in Figure 3, for example, do not belong to the Markov blanket of the figure).

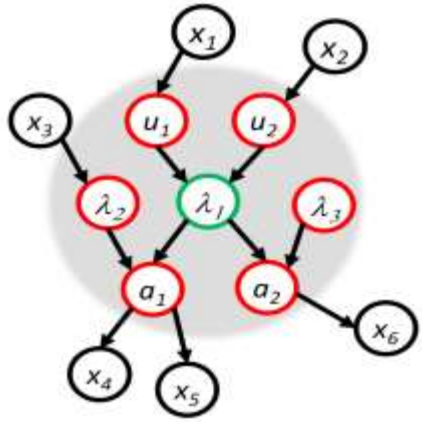

FIGURE 4: Representation of a Markov blanket [Hohwy (2017), p. 3].

Interestingly, Markov blankets have been used as a positive argument for the compatibility between predictive and modular architectures [Hohwy, 2013]. Indeed, one can easily interpret Markov blankets as eliciting some kind of horizontal isolation since the information processed in each Markov blanket may be, in some way, blind to the information processed in other layers of the network. But this argument can be resisted. Modules defined in such a way overlap, and thus, it becomes unclear if the resulting isolation should be considered genuine informational isolation. Remember, however, that Drayson's modularity admits the possibility of overlapping modules and claims that although inconsistent with traditional modularity (a la Fodor), predictive architectures are consistent with the weaker notion of modules that she puts to work. But even ignoring the explanatory loss of a weaker notion of module (which is obvious), the argument fails. Drayson's argument requires minimal functional segregation between nodes, a condition for a module to be considered a module. But if Markov blankets are nested, then instead of an informational closure, what is assembled is a complex network of states dependent on other preceding states in which some kind of informational exchange must be held [see Kiefer (2017), n. 17]. Consider the following illustration: once neurons at n-1 are activated, influenced by the activation at layer n, the generative weights between n and n-1 are established. The subsequent generative weights connecting n-1 and n-2 are engaged by the prior activation at n-1, thus causing a prescribed activation pattern at layer n-2. And once again, once activated n-2, the generative weights between n-2 and n-3 engage, and so forth, down to the layer just above the lowest level of input data, which receives the initial sensory data from the receptors. The point is that this picture fits much better with the idea of functional integration or mean-field factorization than with the idea of functional specialization at every level of the hierarchy. The result is an optimal perspective, where each level in the hierarchy is accountable to the others by determining the accuracy of a perceptual representation and updating this representation accordingly to ultimately provide a consistent representation that resolves uncertainty and reduces redundancy from the external world.

V. Conclusion

Predictive processing reverses conventional views about the flow of information in the brain. Instead of being a mere transducer passively waiting to be activated by external stimuli, perception is seen as an active process that participates jointly with the internally generated model to predict the incoming sensory data by correcting predictions from the actual inputs, thus updating the model accordingly. Far from being anchored to current situations, our brain is constantly predicting what will happen next from previously recorded information. According to PP, perception relies on priors as much as on incoming information.

Attempting to reconcile seemingly conflicting theories is a commendable and sometimes fruitful action. I believe this is not the case here. If the postulates of PP are on the right track, the most we can thank modular architectures is, besides having opened a fertile debate, having distanced us from how minds work. In some way, Fodor overstated modules and, since then, theorists have tried to accommodate them to the new empirical findings. The crux of a module, its informational encapsulation, has been converted over time into functional independence, functional specialization and finally functional segregation, thus losing its original formulation and weakening its explanatory value. Modules cannot admit narrower formulations. This is in itself a problem for modular architectures, but even taking the weakest notion of module, it is not possible to fit it into the increasingly influential theses suggested by predictive architectures. What a module cannot support is that perception and cognition lie inseparably in a continuum (continuity claim) and that no part of perceptual processing is functionally isolated from cognitive processing (non-isolation claim). I have shown, against Drayson, that predictive and modular architectures are opposite, antagonist and mutually excluding. Minds can be seen as functionally decomposed mechanisms assembled by distinct components, but not in the way that modular architectures predict — to be functionally decomposed does not mean to be functionally discontinuous or functionally isolated. Perception and cognition are continuous at all levels, and no part of the perceptual brain is isolated from cognition. Finally, if the predictive picture of the mind is right, modular architectures should be reconceptualized, if not abandoned.

Departamento de Psicología
Universidad Nacional de Educación a Distancia
C/Juan del Rosal 10,
28040, Madrid
E-mail: elsergei@hotmail.com

NOTES

[1] An exception is Sperber (2001), who defends massive modularity without renouncing the informational encapsulation requirement.

[2] The idea of specialized functional modules raises other worries. For example, the specific domain in which a functional specialized module is circumscribed must still be established (for the moment, only arbitrary suggestions have been proposed). Furthermore, even accepting that some modules can interact between them (there is evidence showing the regular transfer of information between distinct sense modalities [see the McGurk effect]), it is not clear the amount of information that can be transferred without violating the criterion of sensory specialization.

[3] Just as modularity relates to vertical, autonomous, non-interactive, localized psychological faculties and to rule-based processing carried out in a step-by-step or serial procedure, connectionism relates to horizontal, interactive, non-modular cognitive systems and to parallel nonlinear dynamic processing [see, however, Drayson (2017), p. 7].

[4] The arguments addressed by Drayson require, and it is critical to recall it, to understand modularity more in the sense of functional independence rather than informationally encapsulated modules; modularity is, in this sense, more a functional than a structural approach.

[5] Note that the layout of the image is crucial, the configuration of the figure is, in fact, decisive for the illusion to work. For example, smaller arrows or longer lines would undo the illusion.

[6] It should be noted that I am not arguing against some sort of functional differentiation in the brain, — perhaps defined by how functions are integrated [Friston (2011)] —, but against the idea that functional differentiation requires modularity. As a matter of fact, functional differentiation is a weaker notion than functional independence or functional specialization. Modularists have moved from positing compacted and encapsulated mental devices that process very specific information, to admitting their porosity on certain occasions, to recognising the persistent influence of cognition over perceptual experience, and from there to define modules as functional independent devices necessary to make computational processes more efficiently and fast. In my view, predictive processing does not need modules in any of these senses, nor as functional independent devices. This is simply because predictive architectures can perfectly make informational restrictions without the need for any modular mechanism [see Parr et al. (2020) and Hipolito et al. (2021)]. For example, the bottom-up/top-down running of information can operate, depending on their influence and prior appearance, as an informational selector, or the degree of prediction error marked by the environmental uncertainty can redirect the information towards *a delimited range of parameters to which their processing is sensitive* [see the functional characterization of Burnston and Cohen (2015), p. 132]. All this is, in my view, sufficient to make the system efficient and fast. Thanks to an anonymous reviewer for pressing me on this point.

[7] Note that cases such as introspective visualizations, episodic memories, mental imagery, dreams or even endogenous hallucinations, which are not perceptual experiences but are felt phenomenologically similar to them, are examples that support the continuity between perception and cognition.

[8] For some researchers sympathetic with the Predictive Processing framework this explanation is insufficient; a more complete explanation would require predictive processing to appeal to bodily and environmental factors [see Gallagher, Hutto and Hipolito (2022)].

[9] For a critical discussion of many of these studies, see Firestone and Scholl (2016).

[10] Eells and Sober (1983) present a theorem in which, seen as a Markov blanket (see the following reason), the weather in the past century can influence today's weather, albeit being an infinitesimal small influence.

[11] Beni (2022) suggests one way to save modularity by claiming that PP can execute DAGs and DCMs as complementary models. Furthermore, Beni argues that even taking DCMs, a modest form of modularity still stands. For the first point, I argue that the dynamics described in DCM fit better with PP's postulates (see below in this paragraph). The problem with the second point is that it requires a very modest form of modularity, a form of modularity that amounts to informational segregation between patterns of connectivity [Beni (2022)]. But, I think this is not enough to avoid transitivity between nodes. In PP, the connectivity patterns are better explained as recursively extended to any arbitrary number of levels giving rise to a functional interconnectivity over long

causal chains, than as specialized and segregated modules [Hipólito and Kirchhoff (2019)]. Undoubtedly, the kind of network executed by brains, just as the appealing to Markov Blankets formalism (see below in this section) opens a great door on the appropriate way of thinking about the architecture of the mind; there is still much to discuss, but the notion of modularity that is intended to be saved seems to me increasingly restricted, and ultimately, explanatorily vacuous [for a very recent special issue on these topics see Hipólito and Kirchhoff (2023)].

REFERENCES

BALCETIS, E. and DUNNING, D. (2010), "Wishful Seeing: More Desired Objects Are Seen as Closer"; *Psychological Science*, 21, pp. 147-52.

BANERJEE, P., CHATTERJEE, P. and SINHA, J. (2012), Is It Light or Dark? Recalling Moral Behaviour Changes Perception of Brightness"; *Psychological Science* 23, pp. 407-409.

BARRETT, H.B. and KURZBAN, R. (2006), "Modularity in Cognition: Framing the Debate; *Psychological review*, 113(3), pp. 628-647.

BECHTEL, W. and ABRAHAMSEN, A. (1991), *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*; Oxford: Blackwell.

BENI, M. D. (2022), "A Tale of Two Architectures: Free Energy, Its Models, and Modularity; *Consciousness and Cognition*, *98*, 103257.

BLOCK, N. (2023), *The Border Between Seeing and Thinking*; Oxford University Press.

BURNSTON, D. and COHEN, J. (2015), "Perceptual Integration, Modularity, and Cognitive Penetration" In A. Raftopoulos and J. Zeimbekis (Eds.) *Cognitive Influences on Perception: Implications for Philosophy of Mind, Epistemology, and Philosophy of Action*. Oxford University Press.

CARRUTHERS, P. (2006), *The Architecture of the Mind*, Oxford: Oxford University Press.

CLARK, A. (2013), "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science"; *Behavioral and Brain Sciences* 36(3), pp. 181-204.

—— *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*; New York: Oxford University Press.

COLTHEART, M. (1999), "Modularity and Cognition. Trends in Cognitive Science"; 3, pp. 115-20.

DEWHURST, J. (2017), "Folk Psychology and the Bayesian Brain"; in T. Metzinger and W. Wiese (Eds.) *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.

DRAYSON, Z. (2017), "Modularity and the Predictive Mind" In T. Metzinger and W. Wiese (Eds.) *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.

EELLS, E. and SOBER. E. (1983), "Probabilistic Causality and the Question of Transitivity"; *Philosophy of Science,* 50, pp. 35-57.

FIRESTONE, C and SCHOLL. B. J. (2016), "Cognition does not Affect Perception: Evaluating the Evidence for 'Top-Down' Effects"; *Behavioural and Brain Sciences*, 39, pp. 1-77.

FODOR, J. A. (1983), *The Modularity of Mind*; Cambridge, MA: MIT Press.

—— *The Mind Doesn't Work That Way*, Cambridge, MA: MIT Press.

FODOR, J. A. and PYLYSHYN. Z. W. (1988), "Connectionism and Cognitive Architecture: A Critical Analysis"; *Cognition,* 28, pp. 3-71.

FRISTON, K. J. (2010), The Free-Energy Principle: A Unified Brain Theory?; *Nature Reviews Neuroscience,* 11(2), pp. 127–38.

—— (2011), "Functional and Effective Connectivity: a Review"; *Brain Connectivity*, 1(1), pp. 13-36.

FRISTON, K. J. (2013), "Life As We Know It"; *Journal of The Royal Society: Interface*, 10 (86), 20130475.

FRISTON, K. J. and PRICE. C.J. (2001), "Dynamic Representations and Generative Models of Brain Function"; *Brain Research Bulletin*, 54, pp. 275–285.

GALLAGHER, S., HUTTO, D. and HIPÓLITO, I. (2022), "Predictive Processing and Some Disillusions about Illusions"; *Review of Philosophy and Psychology* 13, pp. 999-1017. *https://doi.org/10.1007/s13164-021-00588-9*

GREEN, E. J. (2023), "The Perception-Cognition Border: Architecture or Format?"; in B. P. McLaughlin and J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*, Oxford: Blackwell.

HARBER, K. D., YEUNG, D. and IACOVELLI, A. (2011), "Psychosocial Resources, Threat, and the Perception of Distance and Height: Support for the Resources and Perception Model"; *Emotion* 11, pp. 1080-90.

HIPÓLITO, I., and KIRCHHOFF, M.D. (2019), "The Predictive Brain: A Modular View of Brain and Cognitive Function?"; *Preprints.org*, 2019110111.

—— (2023), "Breaking Boundaries: The Bayesian Brain Hypothesis for Perception and Prediction"; *Consciousness and Cognition*, 111, 103510.

HIPÓLITO, I., RAMSTEAD, M. J., CONVERTINO, L., BHAT, A., FRISTON, K., and PARR, T. (2021), "Markov Blankets in the Brain"; *Neuroscience & Biobehavioral Reviews*, *125*, pp. 88-97.

HOHWY, J. (2013), *The Predictive Mind*; Oxford: Oxford University Press.

—— (2017), "Priors in Perception: Top-Down Modulation, Bayesian Perceptual Learning Rate, and Prediction Error Minimization"; *Consciousness and Cognition*, 47, pp. 75-85.

KIEFER, A. (2017), "Literal Perceptual Inference"; in T. Metzinger and W. Wiese (Eds.) *Philosophy of Predictive Processing*; Frankfurt am Main: MIND-Group.

LEE, T.S., and MUMFORD, D. (2003), "Hierarchical Bayesian Inference in the Visual Cortex"; *Journal of Optical Society of America*, 20(7), 1434-1448.

LEVIN, D. T. and BANAJI, M. R. (2006), "Distortions in the Perceived Lightness of Faces: The Role of Race Categories"; *Journal of Experimental Psychology*: General 135, pp. 501-12.

LUPYAN, G. (2015), "Cognitive Penetrability of Perception in the Age of Prediction: Predictive Systems Are Penetrable Systems"; *Review of Philosophy and Psychology*, 6, pp. 547-569.

MACPHERSON, F. (2017), "The Relationship Between Cognitive Penetration and Predictive Coding"; *Consciousness and Cognition*, 47, pp. 6-16.

MCCAULEY, R.N., and HENRICH. J. (2006), "Susceptibility to the Muller-Lyer Illusion, Theory-Neutral Observation, and the Diachronic Penetrability of the Visual Input System"; *Philosophical Psychology,* 19(1), pp. 79-101.

MCCLELLAND, J. L. (2013), "Integrating Probabilistic Models of Perception and Interactive Neural Networks: A Historical and Tutorial Review"; *Frontiers in Psychology*, 4, p. 503.

OGILVIE, R., and CARRUTHERS, P. (2016), "Opening Up Vision: The Case Against Encapsulation"; *Review of Philosophy and Psychology, 7*, pp. 721–742.

PANICHELLO, M. F., CHEUNG, O., and BAR, M. (2013), "Predictive Feedback and Conscious Visual Experience"; *Frontiers in Psychology*, 3, pp. 620.

PRINZ, J. (2006), "Is the Mind Really Modular?"; in R. Stainton (Ed.). *Contemporary Debates in Cognitive Science* (pp. 22-36) Oxford: Blackwell.

PYLYSHYN, Z. W. (1999), "Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception", *Behavioural and Brain Sciences*, 22, PP. 341-365.

QUILTY-DUNN, J. (2020), "Perceptual Pluralism"; *Noûs* 54(4), pp. 807-838.

RAFTOPOULOS, A. (2009), *Cognition and Perception: How Do Psychology and Neural Science Inform Philosophy?* The MIT Press.

RAO, R.P.N. and BALLARD, D.H. (1999), "Predictive Coding in The Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-Field Effects"; *Nature Neuroscience,* 2 (1), pp. 79-87.

ROCHE, W. A. (2012), "A Weaker Condition for Transitivity in Probabilistic Support"; *European Journal for Philosophy of Science*, 2 (1), pp. 11-118

RUDEL, R.G. and TEUBER, H. L. (1963), "Decrement of Visual and Haptic Müller-Lyer Illusion on Repeated Trials: A Study of Crossmodal Transfer"; *Quarterly Journal of Experimental Psychology*, 15(2), pp. 125-131.

SHOGENJI, T. (2003), "A Condition for Transitivity in Probabilistic Support"; *British Journal for the Philosophy of Science*, 54, pp. 613-616.

SIEGEL, S. (2012), "Cognitive Penetrability and Perceptual Justification"; *Noûs* 46, pp. 201-22.

SIMS, A. (2016), "A Problem of Scope for the Free Energy Principle As a Theory of Cognition"; *Philosophical Psychology*, 29, pp. 967-980.

—— (2017), "The Problems with Prediction. The Dark Room Problem and The Scope Dispute"; In T. Metzinger and W. Wiese (Eds.), *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.

SNIJDERS, T. (2012), *Transitivity and Triads*; University of Oxford. Available at *http://www.stats.ox.ac.uk/~snijders/Trans_Triads_ha.pdf.*

SPERBER, D. (2001), "Defending Massive Modularity"; in E. Dupoux (ed.) *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*, Cambridge, MA: MIT Press.

SPOHN, W. (2009), "The Difficulties with Indirect Causation"; *Causation, Coherence and Concepts: A Collection of Essays* (pp. 57–65), Springer.

STEFANUCCI, J. K. and GEUSS, M. N. (2009), "Big People, Little World: The Body Influences Size Perception"; *Perception*, 38, pp. 1782–95.

STERZER, P., FRITH, C. & PETROVIC, P. (2008), "Believing Is Seeing: Expectations Alter Visual Awareness"; *Current Biology*, 18(16): R697–98.

STOKES, D. AND BERGERON, V. (2015), "Modular Architectures and Informational Encapsulation: A Dilemma"; *European Journal for Philosophy of Science*, Vol. 5 (3), pp. 315-338.