Mansell, W. 2022. An Integrative Control Theory Perspective on Consciousness. Psychological Review.
https://doi.org/10.1037/rev0000384

# An Integrative Control Theory Perspective on Consciousness

*Warren Mansell*

*Curtin University, Perth, Australia*

RUNNING HEAD: Control Theory and Consciousness

Correspondence concerning this article should be addressed to:

Warren Mansell

Professor of Mental Health

School of Population Health

Curtin University, Perth, Australia

Email: warren.mansell@curtin.edu.au

## Author Note

Thank you to Liam Mason, Vyv Huddy, Martin Taylor, Eva de Hullu, Tom Scholte, Eetu Pikkarainen, Bruce Nevin, Rupert Young, Brian D'Agostino, Henry Yin, James Kelly, Luke Clark, Kent McClelland and Phil Barnard for their comments and suggestions regarding earlier versions of this manuscript. Thank you to Andrew Bell for graphic design of the figures and the the reviewers of an earlier version of this manuscript.

*Abstract*

An integrative account of consciousness should have a number of properties. It should build upon a framework of non-conscious behavior in order to explain how and why consciousness contributes to, and addresses the limitations of, non-conscious processes. It should also encompass the primary (phenomenal), secondary (access) and tertiary (self-awareness) aspects of consciousness. A number of accounts have proposed a role for consciousness in the prediction of sensory input, yet these proposals do not address how organisms deal with multiple, unpredictable, disturbances to maintain control. According to Perceptual Control Theory (PCT), purposiveness is the control of hierarchically organized perceptual variables via changes in output that counteract disturbances which would otherwise increase error between the current value and the reference value (goal state) of each perceptual variable. In PCT, reorganization is the process required for the adaptive modification of control systems in order to reduce the error in intrinsic systems that control essential, largely physiological, variables. The current article proposes that primary consciousness emerges from this system, and is sustained as secondary consciousness through a number of processes including the control of the integration rate of novel information via exploratory behavior, attention, imagination, and altering the mutation rate of reorganization. Tertiary consciousness arises when internally sustained perceptual information is associated with specific symbols that form a parallel, propositional system for the use of language, logic, and other symbolic systems. The hypotheses and initial research designs to test this account are provided.

Keywords: consciousness, awareness, self-regulation, cybernetics, prediction, perceptual control theory

## Scientific Background

The science of consciousness is a critically important field that intersects psychology, psychiatry, neuroscience and philosophy. There is no consensus on its definition, yet three levels are often delineated (Block, 1995; Montemayor & Haladjian 2015; Vandekerckhove, Bulnes, & Panksepp, 2014; Vandekerckhove & Panksepp, 2009; see Table 1). Primary, anoetic, or phenomenal consciousness describes a basic state of the affective, homeostatic, and sensory-perceptual experiences that emerge as *qualia* - the fundamental units of experience (see Vandekerckhove et al., 2014). Secondary, noetic or access consciousness describes the capacity to maintain units of experience where they can be associated together, operated upon, reported, and potentially shared across widely distributed systems (e.g. Koriat, 2000). Tertiary, autonoetic consciousness, or self-awareness, describes the largely human capacity to mentally simulate hypothetical experiences within an imagined past or future, and it is often considered alongside the capacity for symbolic systems, tool use and culture (e.g. Kotchoubey,, 2018).

A recent systematic review concluded that no current theory of consciousness accounted for all of the findings, and that an integrative 'cross-talk' between the theories is needed (Yaron, Melloni, Pitts, & Mudrik, 2021). An integrative theory of consciousness would need to explain all three recognized aspects of consciousness. Arguably, it would also need to emerge from a model of non-conscious behavior that has specific limitations that the emergence of consciousness would effectively address. Two of the most widely discussed theories of consciousness are Global Workspace Theory (GWT; Newman & Baars, 1993) and Integrated Information Theory (IIT; Tononi, 2008; Tononi, Holy, Massimini, & Koch, 2016).

GWT proposes a dynamic workspace for bringing together multi-modal perceptual inputs and selecting appropriate outputs. When information enters this workspace, it is conscious. The GWT provides a 'theatre of the mind' that approximates well with people's reports of their own conscious experience: object and agents occurring together with their own bodily perception, thoughts and feelings within an integrated scene that is organised spatially. For example, people can typically report that thoughts occur ‘at the back’ of their mind, or that visual images are superimposed on the visual scene in front of them, and there is a moveable spotlight of attention focused on specific elements. The GWT is consistent with the limited capacity, sequential nature of consciousness; it is specified to the extent that it can be tested using computational models (Baars & Franklin, 2009), and an empirically grounded variant of the model has been specified within the domain of neuroscience (Dehaene & Naccache, 2001).

*Table 1. Overlapping terms within the existing literature that describe three aspects of consciousness*

| | Level | | |
|---|---|---|---|
| Framework | 1 | 2 | 3 |
| Tulving (1995) | Anoetic | Noetic | Autonoetic |
| Block (1995); Montemayor & Haladjian (2015) | Phenomenal | Access | Self-Awareness |
| Vandekerckhove & Panksepp, 2009) | Primary process | Secondary process | Tertiary process |

A key critique of GWT is that it only explains secondary or access consciousness (See Table 1) - and does not address the ‘hard’ problem of consciousness, namely how conscious ‘qualia’ arise from the physical matter of the brain (Chalmers, 2007). Conversely, IIT has contributed a solution: qualia are units of information that cannot be reduced to any smaller units. According to IIT, consciousness can be indexed by phi, which is an observer-independent measure of the optimal integration (or simplification) within any cause-and-effect grouping of elements (such as neurones within the brain) that maximises its informational content. This grouping of elements completely describes both the quantity and quality of experience because the information is utilised for the agent itself, rather than for an external observer to interpret and utilise this information (such as when a human user makes use of the information processed by their smartphone). In support of IIT, measures of phi have shown a relationship with some established indications of consciousness (Kim et al., 2018). However, IIT has been critiqued for a variety of reasons, for example that its causal structure deems it false or outside empirical science (Doerig, Schurger, Hess, & Herzog, 2019) or that IIT would entail that non-living, non-purposeful machines would be conscious (Cerullo, 2015). This latter limitation leaves open the questions of how processes, or living organisms, that are not conscious might function, and why information integration exists - "what purpose does consciousness serve?" - both from an organism's point of view, and from an evolutionary point of view. Arguably, similar criticisms apply to other purely mathematical models of consciousness (e.g. Signorelli, Wang, & Khan, 2021) that omit biology and psychology from their explanations.

There are a wide range of theoretical accounts that attempt to answer the "Why?" of consciousness through the answer that living organisms need to predict their sensory inputs in order to survive and reproduce (e.g. Deane, 2021; Rudrauf, Bennequin, Granic, Landini, Friston, & Williford, 2017; Safron, 2020; Seth, Suzuki, & Critchley, 2012; Seth, 2021). The starting point for these accounts is within homeostasis: Living organisms must keep essential physiological variables (e.g. blood glucose levels, salt ion concentrations) within bounds that allow them to stay healthy and survive. Yet, in order to pre-empt going outside these physiological limits before it occurs (allostasis), the organism needs to make predictions about itself and its surroundings (e.g. Pezzulo, Rigoli, & Friston, 2015). This is proposed to be in the form of an internal model of itself and the world that it optimizes by testing its beliefs against sensory inputs from its body and the environment. Within predictive models, the discrepancy between the expected input and actual input is the prediction error that denotes consciousness, and so this account goes some way to explain 'why' phi may exist - as an index of prediction error that is used for *epistemological foraging* - acquiring information to update and optimize existing predictive models of the self and world (Köster, Kayhan, Langeloh, & Hoehl, 2020). Indeed, a similar concept of *innovation rate* or 'rate of cognitive structural change' was introduced by McReynolds (1971) in the context of information processing. Predictive accounts tend to rely upon generating and updating probabilistic inferences regarding the likelihood of future rewards, including predicting the perceptual consequences of potential actions. There are a number of key learning parameters within predictive models, such as learning rate and inverse temperature (Schweighofer & Doya, 2003), that are likely to be involved in this construct. As stated by Layzer (1975): "The present moment always contains an element of genuine novelty; and the future is never wholly predictable" (p69, Layzer, 1975).

A number of commentators have nonetheless pointed to the limitations of predictive processing accounts of consciousness (e.g. Marvan & Havlík, 2020; Schlicht & Dolega, 2021). For example, the multiplicity of theories in this field are each based on different assumptions about the nature of the predictive processing architecture. These accounts also assume that prediction is the fundamental process in the brain, whilst delegating the process of control itself to either homeostatic systems, or to specific outputs that maximize reward or 'value'. Yet, the social and natural environment encompasses multiple disturbances that need to be counteracted for an action to have its predicted effect in the world. Some disturbances, such as an uneven surface or turbulence, are not predictable. Others are predictable, but would require high amounts of computation and high amounts of data to make useful predictions, such as predicting the paths of people rushing through a crowded train station in order to avoid them. This issue is particularly stark in the arena of robotics, in which devices driven by predictive models consistently require negative feedback control systems to function effectively (Argentim et al., 2013; Johnson et al., 2020; Vannucci et al., 2015). This calls into question whether prediction itself can be used as a basis for an account of consciousness, and raises the possibility that negative feedback control may be the core principle to construct an account. Indeed, predictive processing models of consciousness require a threshold at which the purposeful functioning of an individual would arise from predictive processes. Yet what if purpose is intrinsic and precedes consciousness? In answer to this, Perceptual Control Theory (PCT; Powers, 1973; Powers, Clark & McFarland, 1960a,b) has its basis in negative feedback control and it provides a novel perspective on consciousness that does not rely on prediction. In this article, I will also explain how PCT provides a functional model of the non-conscious, yet purposive, behavior of living organisms upon which consciousness develops, and how its architecture allows the modelling of all three - primary, secondary, and tertiary - aspects of consciousness.

PCT explains how an agent can be agile to deal with multiple, dynamic disturbances that cannot be predicted ahead of time, and so would not be addressed by predictive computational models (Barter & Yin, 2021; Matic, Valerjev, & Gomez-Marin, 2021; Young, 2017). Based on negative feedback control, PCT uses a simpler and more parsimonious architecture than required for predictive modelling (Barter & Yin, 2021; Matic & Gomez-Marin, 2021; Young, 2017). Importantly, the main tenets of PCT have received wide empirical support (for a review, see Mansell, 2020). The PCT approach can be tested through constructing PCT computational models and evaluating their performance against data collected from the real world. For example, a systematic review of manual tracking found a consistently robust match between PCT computational models and human movement data (Parker, Willett, Tyson, Weightman, & Mansell, 2020). A series of studies involving PCT models that control perceptual variables - retinal linear velocity and lateral displacement - closely match real-world movement data during object interception without requiring prediction (Marken, 2021). Furthermore, software that implements perceptual control systems in robots permits effective navigation and obstacle avoidance (Young, 2017), balance (Johnson et al., 2020), locomotion on four legs (Barter & Yin, 2021), and arm coordination (Matic & Gomez-Marin, 2022). These robots exhibit simple, naturalistic, purposive behavior through the real-time control of perceptual input, without the need for prediction, or consciousness. Therefore, the PCT architecture provides a basis of non-conscious behavior from which to explain how consciousness might emerge to resolve its limitations and enhance its functionality. This foundational level of non-conscious, purposive behavior will be explained first.

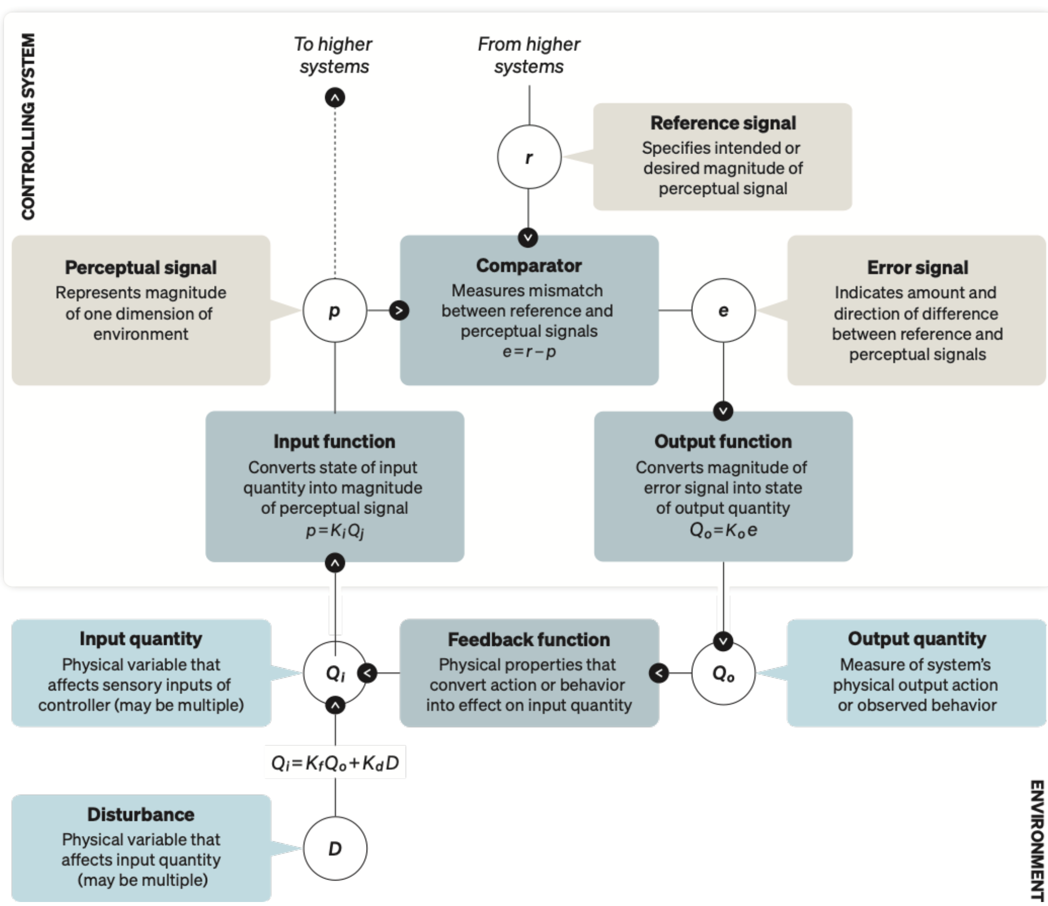
## Purpose as the Hierarchical Control of Input Underpins the Development of Consciousness

### *Control of Input*

Like the predictive processing accounts, PCT is also grounded in an understanding of the control processes that occur within homeostasis and allostasis. Put simply, control is essential for life (Cziko, 2000; Noble, 2008). According to PCT, behavior is a part of the body's systems that control input. Control of input - rather than prediction - is the essence of the purposiveness of living organisms, and it is proposed to be in-built (Powers, 1978). Control of input is necessary to survive in a world of multiple, unpredictable disturbances, and therefore the purposive functioning of a living organism precedes the development of consciousness according to PCT. So, an account of consciousness is not required to explain how purpose arises (Carey, 2018; Pikkarainen, 2021). But how does control of input work?

According to Powers, control is "the achievement and maintenance of a preselected perceptual state in the controlling system, through actions on the environment that also cancel out the effects of disturbances" (p283; Powers, 1973). PCT specifies a closed loop between the organism and its environment such that discrepancies (error) between the current value of a perceptual input and its reference value is continuously, and automatically, corrected, in order to maintain control (see Figure 1). This process relies upon the construction of a *perceptual signal* within the nervous system by an *input function* that integrates and transforms input signals that ultimately have their origins within the diverse sensory organs of the body.

*Figure 1.* *A model of a control unit within PCT. These units are connected together in a branched hierarchy, which is not shown in this diagram, but connections to higher level units are indicated. The diagram illustrates how perceptual inputs are controlled without the requirement of consciousness within a closed loop between the organism and its environment. Perceptual signals representing the magnitude of a variable are constructed by an input function which combines inputs from lower-level systems that ultimately derive from sensory signals generated from aspects of the environment. The unit compares the magnitude of the perceptual signal with a reference value for the signal to generate an error signal which is transformed by an output function to set the reference values for lower-level systems that ultimately interface with the body and environment to transform its effects, via a feedback function, to counteract the effects of disturbances to the input quantity being controlled. These units are necessary for consciousness but they operate outside conscious awareness. Source: original diagram by William T. Powers redrawn by Dag Forssell (2016, p. 32); reproduced with permission.*



Within PCT, there is no explicit role for 'prediction' within the architecture. As stated by William T. ,, "There is no prediction of the future involved; the organism simply acts at all times in the direction that will oppose the present-time error." (p. 230, Powers, 1980). Instead PCT clearly separates the *specification* of a perceptual variable by an input function, from its *control* through a closed loop with the environment. The perceptual signal is internal to the nervous system, but it is not part of a predictive model. Rather, it is a continuous variable that covaries with an aspect of the self and world. Yet the selection of the most appropriate perceptual variable for control is based on whether controlling it in turn facilitates control by higher level systems. For example, the perception of current velocity at one level can be controlled in order to control perceived distance from a target at a higher level (Johnson et al., 2020). The variable need not have a *direct* correspondence with a real-world physical measurement, need not make predictions, and may even be constructed purely from the agent's own egocentric perspective (Rabinovich & Jennings, 2010). A PCT system nonetheless generates all of the features described by other accounts as 'prediction', but it separates out the potentially ambiguous meanings of this term into the specification of a variable, and the control of a variable. In the absence of observable action, perceptual control may look like 'prediction', but within PCT, this is simply where either (a) control requires no action (such as when error is zero); (b) action is prevented (such as in many animal experiments); (c) where another control system is preventing action (such as when action would have undesirable consequences), or (d) where internally simulated perception provides a substitute for the input generated by action (such as during imagination; see later).

*Hierarchies*

Powers specified types of perceptual variables and organized them within a hierarchy of levels (Powers et al., 1960a,b; Powers, 1973), which are envisaged as a branching network, rather like a root system (Goekoop & de Kleijn, 2021); see Figures 2 and 3. The lower levels specify continuous perceptions of the body and the environment, such as the *intensity* of light, pain, or pressure, and increasingly higher levels specify increasingly more abstract perceptions formed from organizations of inputs from levels below, culminating in *principles* (e.g. honesty), and then *system concepts* (e.g. self identity), at the highest level. The hierarchy works automatically to control perceptual variables at each level. For example, if a woman sees herself as a good person (*system concept*), she would uphold the *principle* of honesty (among other principles), and she would engage in activities that are consistent with this principle. These activities would rely upon output signals that go down to the level below and specify the reference values for a perceptual variables controlled at that level. For example, if the man walking in front of her dropped his wallet, she would pick up the wallet and give it to him. This is a *program* that would specify as its output to the next level down a *sequence* of perceptions: One possible sequence would be: woman picks up the wallet; calls to the man; the man sees the wallet; the woman gives man the wallet. At further levels down the hierarchy, the output signals form the reference values for a cascade of controlled perceptions in order to be realized in action as the muscles of the body affect physical aspects of the world that are continuously sensed and fed upwards as input signals. Anecdotally, it appears that an individual is only consciously focused on one controlled variable, at one level, at any point in time, and that this focus shifts over time: 'be an honest person' - 'pick up the wallet' - 'give it back' . And yet the activity runs continuously, largely outside awareness. So, what is the reason for the shift in the focus of conscious awareness? The answer lays within an understanding of reorganization, and conflict, discussed next.

*Figure 2. Illustration of the connections between components in a perceptual hierarchy. The boxes are functions and the arrows show the passage of signals. Each control unit is simplified as an Input Function (I), which sends an input signal to a Comparator (C) that subtracts it from a reference signal, to generate an output signal (O). The top level of the diagram shows how the output signals from each level are sent downward and converge (via a reference input function, not illustrated) to form a reference signal for each comparator at the next level down. The second level of the diagram shows how a copy of each input signal is sent to a higher level where they converge on an input function to generate the input signal in a higher-level unit. The bottom level shows the combined input and output connections as in a working hierarchy. For simplicity the connections for the outer units are not shown.*

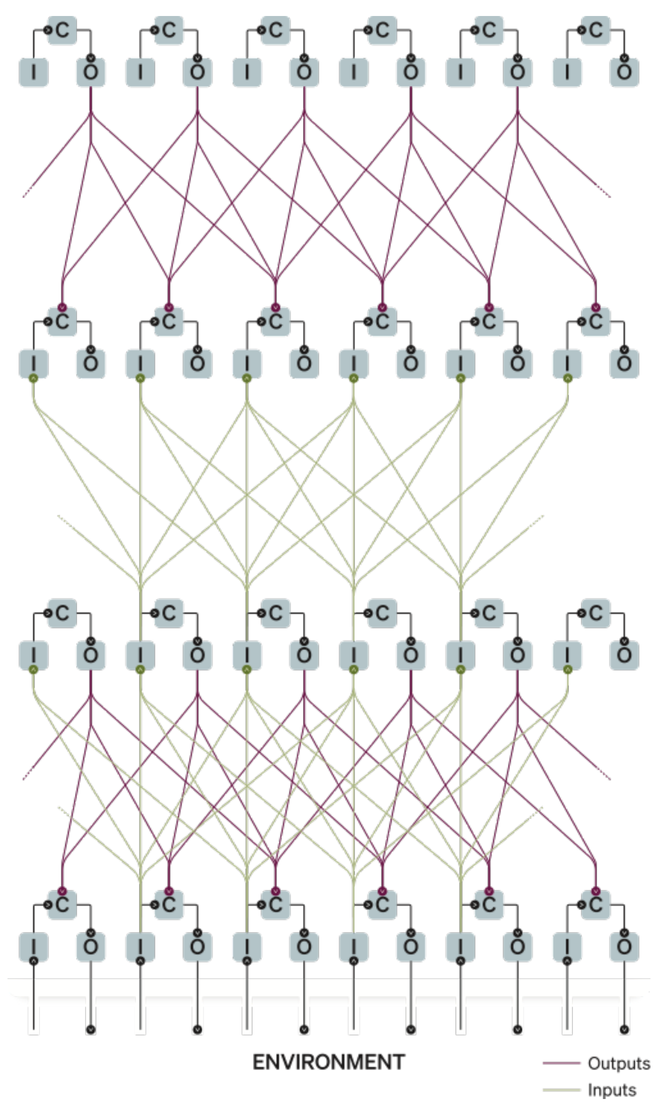
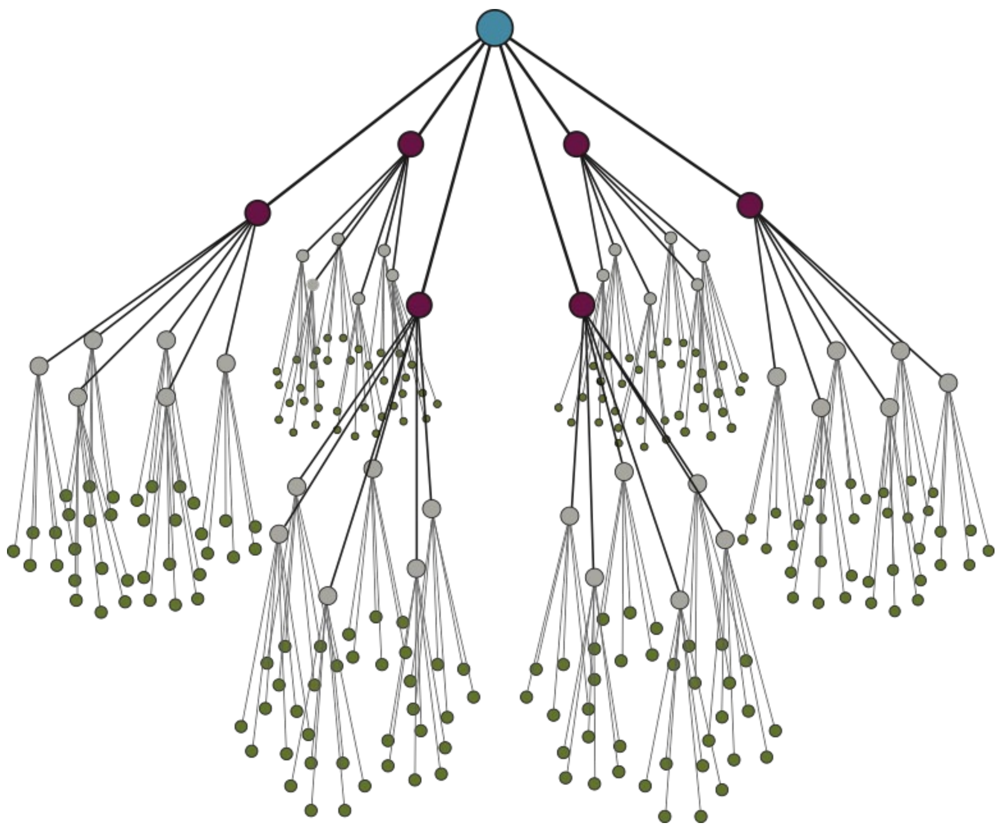


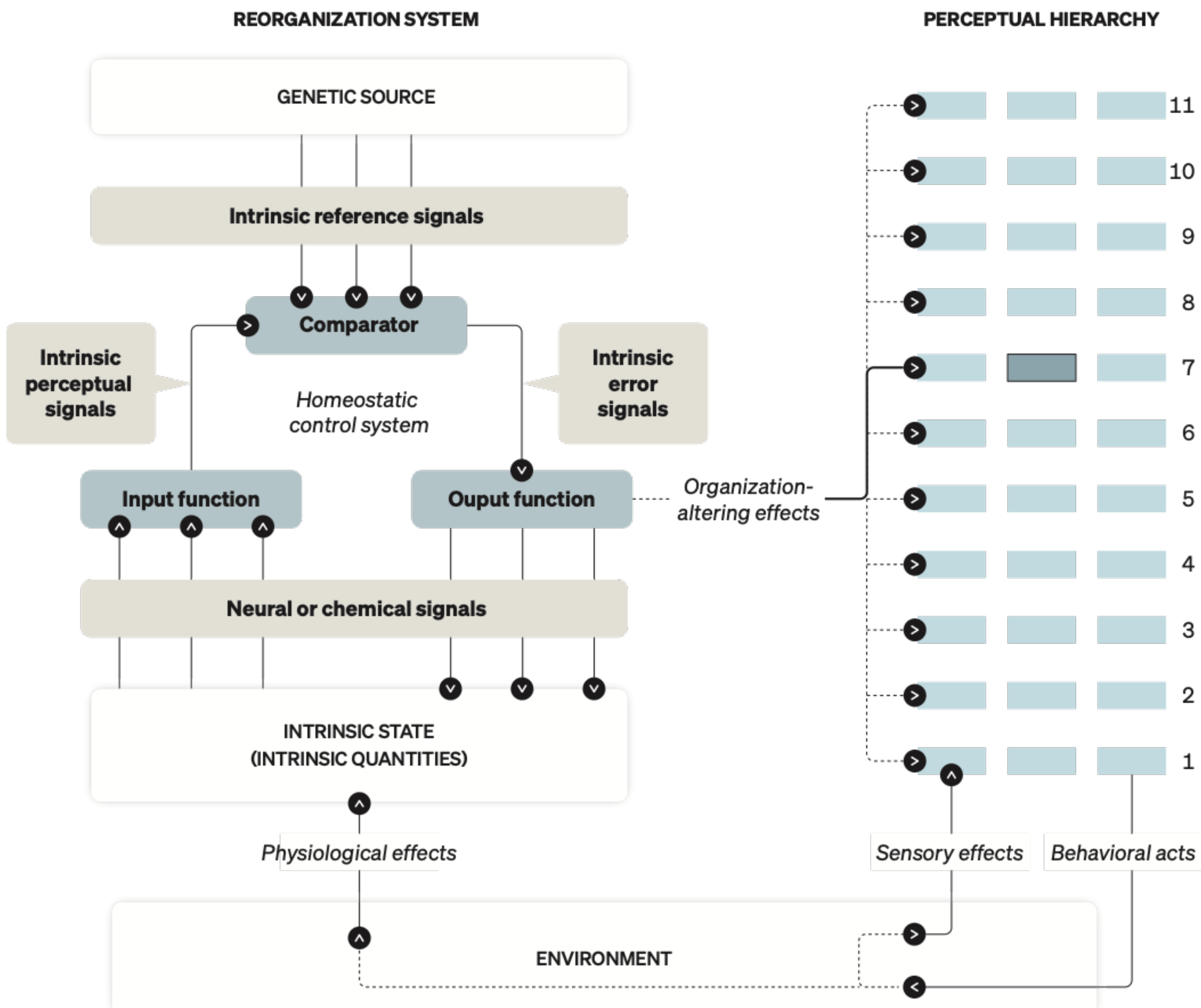
*Figure 3.* *This branching root diagram illustrates the likely structure of the connections such that multiple lower-level units connect to the level above. For simplicity, this diagram does not show the functions or pathways of signals, and is only a sub-branch of a larger hierarchical system.*



*Reorganization*

To enable learning, Powers needed to specify an in-built process responsible for constructing and modifying control units throughout the hierarchy. Reorganization is the learning algorithm posited by Powers, based in part on Ross Ashby's notion of requisite variety (Powers, 2008), and on the achievement of negentropy within cybernetics and information theory (p. 79; Powers et al., 1960a). Powers proposed that reorganization is driven by intrinsic error. Intrinsic error occurs when biologically prepared - intrinsic - control systems deviate from their reference states; these are the same homeostatic or allostatic systems governing essential variables as described earlier. The intrinsic systems may govern a range of physiological variables (e.g. blood glucose) and sensory variables that have supported survival throughout evolutionary history (e.g. the smell of nutritious food), although their exact architecture is not specified within PCT. Importantly, whilst trial-and-error learning is a common feature of other learning theories (cf. Shah, 2012), the PCT perspective clarifies that it is the properties of perceptual control systems that are reorganized rather than any specific behavior being reinforced, because as demonstrated earlier, action varies dynamically to counteract current disturbances. When reorganization occurs, randomly generated changes are made to the input functions, output functions, and the parameters (e.g. gain, smoothing constants) of that unit until error reduces. Reorganization therefore emulates natural selection of variants in order to optimize control (Czkio, 1997). Consistent with the PCT account, there is evidence of neural 'reorganization' during uncontrollable stressors that dysregulate the hypothalamic-pituitary-adrenergic (HPA) axis (Huether, 1999). Yet, within the PCT approach, reorganization facilitates learning throughout a continuum of intrinsic error states to reduce them to zero, rather than solely during 'uncontrollable' stress.

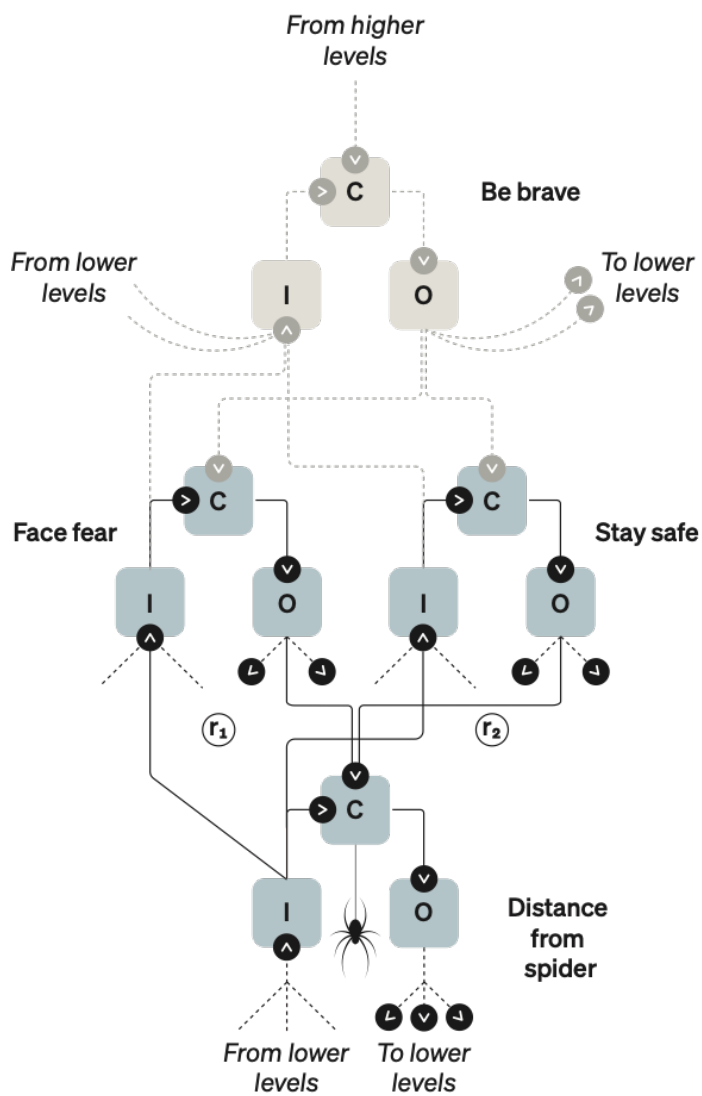
*Figure 4.* *The relationship between the reorganization system and the (learned) perceptual control hierarchy. Intrinsic error drives trial-and-error changes in the properties of the perceptual control hierarchy at a specific location (the arrow leading to a box shaded grey). For simplicity, the connections between control units within the learned hierarchy are not shown. Diagram from Forssell (2016, p. 106) with permission.*



*Conflict*

Reorganization is required to reduce the enduring error that emerges from conflict (p316, Powers et al., 1960b). Conflict occurs when two or more control systems that specify differing reference values compete to control the same perceptual variable. For example, a person who is afraid of spiders, but wants to overcome their fear, is in conflict regarding how close they want to be from a spider at any one moment. Their behavior often oscillates between approach as they attempt to face their fears, and avoidance when the spider does something (e.g. moves suddenly) that prompts a protective action. The fact that it is perceived distance rather than actual distance that is controlled is clear because the same pattern is observed within virtual scenarios on a computer (Healey, Mansell, & Tai, 2017). Conflict can occur at all levels in the perceptual hierarchy, and computer simulations reveal that even skilled movement requires a superordinate control system to organize the controlled inputs of opposing muscles (p. 127; Powers, 2008). Whilst a certain degree of conflict is manageable, enduring conflict can lead to loss of control and needs to be resolved to re-establish control. Within PCT this is achieved through reorganization of the superordinate system in the hierarchy. Figure 5 illustrates how reorganization of a higher-level control unit may resolve conflict in the spider phobia example. Conflict can be resolved by controlling a perceptual signal at a higher level in the hierarchy - for example by controlling for ‘bravery’ - the principle that one can face one’s fears despite the challenges. Indeed, any multi-componential control system is at risk of functionally separate systems controlling variables that are also being controlled by other parts of the same system. This is compartmentalization, which is adaptive for the focused pursuit of a goal, but it can maintain conflict and entail loss of control in the form of dissociations, unless the components are reorganized (Mansell & Carey, 2013). Thus, conflict is a consequence of a non-conscious, hierarchical control system, and the integrative model presented here will explain how it is a problem that can be resolved in increasingly sophisticated ways through the primary, secondary and tertiary aspects of consciousness.

*Figure 5. A simplified diagram of conflict in which each control unit is represented by an input function (I), a comparator (C), and an output function (0). In this example, the controlled variable is the distance from a spider, which has opposing reference signals set by separate superordinate control systems: a low value set by the control of a perception of the self as facing one's fears (r1) and a high value set by keeping safe (r2). Many individuals oscillate between these two distances, leading often to lack of progress with the phobia. In this example, a higher level system controlling the perception of the self as 'brave', is able to balance the two conflicting systems. Original illustration.*



*Automatic Purposiveness*

Each of the features described so far - closed loop control of input, hierarchies, reorganization, and the resolution of conflict - do not entail conscious experience *per se*; indeed they are processes that can be quantified and modelled (Powers, 2008). The fact that a working control hierarchy does not require consciousness to function is consistent with the evidence that even complex behavior, ostensibly used to meet high level goals such as routines, and personal principles, appears to be managed outside consciousness (Bargh, Schwader, Hailey, Dyer, & Boothby, 2012, although for a methodological critique, see Vadillo, Konstantinidis, & Shanks, 2016). A classic example is the complex activity that many people manage whilst asleep including, for example: dancing, searching for treasure, greeting and giving lessons (Oudiette, De Cock, Lavault, Leu, Vidailhet, & Arnulf, 2009). Despite the capabilities possible outside consciousness, a control hierarchy provides the foundation for automatic, complex and purposive action that I propose is a *necessary precondition* in order to precisely identify what role consciousness itself serves.

## Primary Consciousness: The Focused Reorganization of a Perceptual Hierarchy by Competing Intrinsic Systems

*Focused Reorganization*

Earlier accounts, including those of predictive processing, have emphasized the role of consciousness in conflict resolution (e.g. Clark, 2018; Gray, 1995; Mayr, 2004). According to PCT, conflict is defined by having opposing goals for the same perceptual variable; it is not defined by the accuracy of perception or by the selection of competing actions. Powers recognized a close relationship between conscious awareness and the focus of reorganization. He proposed that consciousness is drawn to a unit with unresolved error within the perceptual hierarchy and the inputs to this unit then become available to conscious awareness (p318-321, Powers et al., 1960b; p200-201, Powers, 1973). As conscious awareness is focused and sustained there, random adjustments result in reorganization of the control systems involved, allowing control to be restored. It is vitally important that living organism targets and prioritizes certain locations for reorganization over others in order to resolve conflicts, and it needs to control this focus.

*Control Information*

The background of Integrated Information Theory (IIT) is relevant at this point within the PCT model. An apparent advance of IIT was to shift from the classic notion of information as observer-dependent (Shannon, 1948) to a notion of intrinsic information that is observer-independent because it is presented by a system to itself via feedback (Lombardi, & López, 2018). Just like IIT, PCT proposes that systems feed back information 'for themselves'. Yet, it explains that this feedback process is an essential feature of life - maintaining control in a disturbance-prone world - and it has its foundations prior to the emergence of consciousness. PCT illustrates how feedback forms a component of a closed-loop control system that is organised in a highly specific manner (see Figure 1).

An alternative view of information sits in between the Shannon definition that delegates the interpretation and use of information to the observer, and that of intrinsic information within IIT that is self-contained and observer-independent; it is *control information* (Corning, 2007). Control information is "the capacity (know how) to control the acquisition, disposition and utilization of matter/energy in purposive (teleonomic) processes" (p302; Corning, 2007). It is the information used within a control system whilst it is in the process of controlling. The use of control information is inherently a process involving feedback, but it is a specific use of that feedback within a control system or hierarchical network of control systems that defines control information. The next step in bridging IIT and PCT is to consider the integration of input signals to form an input function that occurs at each successive level of the perceptual hierarchy. Thus, the control of a specific perceptual variable (e.g. shape of a drawing of a cat) requires an input function that integrates input signals from a lower level (e.g. functions that specify the head, body and tail), which in turn integrate inputs from a lower level (e.g. the eyes, nose and mouth of the head) in order to specify this variable. Whilst such hierarchical models of pattern recognition have long been proposed (e.g. Selfridge, 1959), they omit the role of perception as a vital part of a control system. The PCT hierarchy provides this. As stated earlier, the PCT hierarchy works automatically without consciousness, despite its use of integrated control information. So where is the role of consciousness? I propose that consciousness emerges from the *novel* integration of lower-level input signals during control, particularly when in situations of continuing conflict.*Qualia as Novel Information Integration*

Figure 4 shows the intrinsic systems within an organism condensed into one unit. Yet in real life, Individuals have many intrinsic systems for a wide variety of basic needs - hydration, energy, specific nutrients, warmth, lack of pain, etc - and these may sometimes be in conflict with one another. For example, if an individual can manage to eat even when they are feeling sick through illness, this will raise their chances of survival (Clark, 2018). There needs to be a system for how the focus of reorganization is selected across multiple intrinsic systems. The integrative control theory model places the emergence of qualia, at least on a very 'immediate' basis, at this point. Qualia emerge as information from diverse lower level systems is integrated by an input function to specify a novel, often more abstract, controlled variable that resolves ongoing conflict and re-establishes control. A similar account has been provided to explain how new 'perceptual worlds' emerge during human infancy according to PCT (Plooij, 2020). The creation and reorganization of a superordinate perceptual variable would allow individuals to prioritize, balance and configure a range of lower-level controlled variables. Returning to the spider phobia examples in Figure 5, primary consciousness occurs during conflict when input signals converge to form a perception of 'bravery' that emerges for the first time within an individual - it is this felt sense that dominates awareness and is constructed from inputs from lower-level systems that are in conflict.

The integrative model presented here is partly consistent with theories proposing that homeostatic, affective and motivational systems are foundational to the emergence of qualia, upon which secondary and tertiary consciousness develops (e.g. Vandekerckhove et al., 2014). However, the current account proposes that non-conscious (but nonetheless complex) control needs to be understood and modelled first, and that consciousness may emerge within a living organism whilst novel perceptions need to be constructed within the moment as these intrinsic systems conflict with one another and require prioritizing. Yet, if this process were the only source of consciousness, it would be fleeting and fragmented, which is only the case in comparatively rare states of mind in humans (Mansell & Carey, 2013). So what are the processes that account for the *continuity* of consciousness? This is the topic of the next section.

## Secondary Consciousness:

## Control of Information Integration Rate, Memory and Modes .

### *Control of Novel Information Integration Rate*

It has been proposed that reorganization operates like the genetic algorithms inspired by natural selection (Czkio, 1997; Mansell, Carey, & Tai, 2015). Genetic algorithms operate most effectively using an optimum mutation rate (Clune et al., 2008). Thus, from a PCT perspective, the novel integration of input signals would not only emerge from the binding of diverse sources of incoming information but on how these inputs are transformed. Indeed, if the rate of novel information integration were to be an endogenously controlled variable in itself, this might sustain qualia over time. Thus, based on PCT, I propose that living organisms who develop perceptual hierarchies are intrinsically motivated to find new ways to integrate information at a self-governed rate, and they do so 'creatively', testing out different transformations (e.g. different delays, derivatives) to their input signals to continually optimize the specification and control of a perceptual variable. The information integration rate could be controlled either by accessing novel perceptual inputs through actions such as eye movement, object manipulation, locomotion, or social interaction, or otherwise by altering the intrinsic, trial-and-error mutation rate of reorganization that is applied to these inputs. Humans can also use a third method to add variants in input - their imagination, as described later. The novel integration of inputs would be somewhat similar to McReynold's (1971) concept of innovation rate, yet it would be a requirement in the current model that the function of the integration is to establish a distinctive perceptual signal which can then be controlled; it is not to predict the outcomes of specific actions. In fact, many of the features of advanced, and most likely conscious, animals, and especially humans, can be considered as emergent properties of maintaining novel information integration rate as a controlled variable.

*Continuity:* If information integration rate is kept constant over time, this may help account for the typical continuity of consciousness, and also for occasional discontinuities and exceptions. For example, the experience of ‘time slowing down’ may be a side-effect of an exceptionally increased novel information integration rate when intrinsic systems are in high error states (Hancock & Weaver, 2005). This appears to occur in life-threatening instances when raising novel information integration rate would be a priority over maintaining continuity. The discontinuities in consciousness described as various forms of dissociative phenomena are examples of where multiple compartmentalized control systems are operating within the same individual (Mansell & Carey, 2013); continuity in consciousness may be regained only by information integration across these systems through the development of a higher level control system via reorganization.

*Creative activities:* Curiosity, exploration, creativity, and play may emerge as behavioral attempts to maintain novel information integration rate through accessing new combinations of inputs that are yet to be integrated as an abstracted controlled variable. For example, the diverse experiences of play fighting in the young help the individual to learn how to perceive and manage their relationships with others, the rules of various games, and the principles of conduct (Pellis & Pellis, 2017). Within humans, the playful, ‘neotic’ spirit persists into adulthood where it continues to promote the learning of abstracted perceptions: culture, aesthetics, and design for example.

*Avoidance.* If information integration rate is a controlled variable, then there will also be instances in which the individual engages in actions that attempt to reduce the rate of unintegrated perceptual input. Presumably this would be most effectively achieved through behavior such as eye movement away from the source of the unintegrated input, keeping a distance from it, and avoiding environments in which it is encountered. Interestingly, this account converges on the personal accounts, theory and evidence regarding autism (e.g. Frith & Happe, 1994; Iarocci, & McDonald, 2006). Thus, the experiences of a person with a diagnosis of autism may emerge from an intrinsically lower novel information integration rate than other people.

*Memory and Modes*

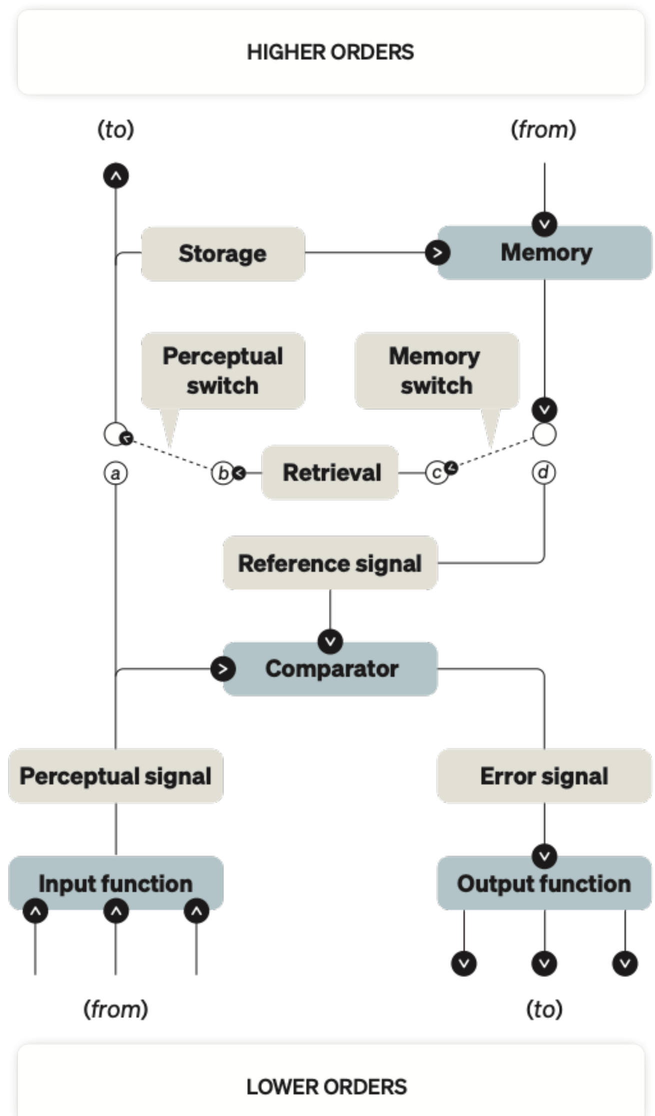
Earlier theorists have identified the role of memory in the development of consciousness (e.g. Tulving 1985). Within PCT, memory is important in order to explain how control systems can function with intermittent perception (e.g. tracking a car moving under a bridge) and also to explain how control systems engage in a variety of modes that form part of conscious experience.

Powers incorporated local memory into each unit of the hierarchy (Powers et al., 1960a; Powers, 1973). Powers proposed that memory allows a control system to operate in four different modes (p. 220; Powers, 1973). These are controlled, automatic, passive observation and imagination modes. The controlled mode involves upward and downward signals between all levels in the hierarchy. The automatic mode allows lower-level systems to operate without receiving reference signals from the level above, but using those stored from memory. Passive observation involves taking in input without sending a signal downwards for control. This is the mode we tend to experience when simply observing a scene, or viewing an object; it is coincidentally the mode often engendered by experimental studies in which a ‘stimulus’ is presented to a participant who is given a behavioral instruction, rather than being allowed to control their input as they would do naturally outside the experiment (Mansell & Huddy, 2018; Marken, 2021). If an individual does not act on their external environment to keep ongoing perceptual input at its reference value, then perceptual error builds up and another process is required to reduce the error. One of these error-reduction solutions is the imagination mode which involves re-routing memory as input ‘as if’ it is being currently experienced, so that higher level systems can receive inputs at their reference value, without engagement with the environment. Clearly, this mode refers to the basis of what is now described as mental simulation (Markman et al., 2012). However, if imagination is insufficient to reduce error then, naturally, reorganization is required to reduce any prolonged error.

Reorganization of a control system during passive observation is a ‘passive’ form of control that is likely to be closest in nature to the processes of cognitive assimilation, or inference, that are described within the predictive processing theories summarized earlier.

Different sections of the control hierarchy can be operating in different modes at the same time – accounting for why we can simultaneously drive a car, plan our route home, and passively observe the scene around us. Figure 6 illustrates how the modes can be varied through four different settings of a ‘switch’ within each control unit. Thus, PCT provides an architecture to reproduce the distinctions between modes of human consciousness, and in particular, how the contents of consciousness may have a single focus (focal consciousness) but can include background perceptions (within ‘fringe’ consciousness; James, 1890) of the body and the environment, alongside memories or imagined perceptions. However, the apparent volitional control of the central focus of consciousness itself is critical to understand, and is the topic of the next section.

*Figure 6. The role of memory and modes in PCT. The switch is shown set in imagination mode, in which memory is recycled as current perception. The four other settings lead to the other modes. In controlled mode, both switches are down and the system functions as if memory was not required. In the automatic mode, the memory switch is set to the reference signal so that memory can be used as a substitute for a reference signal from a higher level system, but the perceptual switch is off so that inputs are not sent upward in the hierarchy for higher level control. In the passive observation mode, the perceptual switch allows the input signal to be sent upward in the hierarchy, but the memory switch is directed away from the reference signal so it is not used for control by lower-level systems.*



According to the control theory perspective presented here so far, the current focus of secondary consciousness is often on the perceptual variable that requires better control through reorganization. When reorganization is reducing error successfully, this account is somewhat analogous to the flow experience (Csikszentmihalyi & Csikszentmihalyi, 1992). Reorganization needs to be sustained on a specific variable in order to have sufficient time to 'stumble' upon a new organization of inputs or other parameters that reduce error and enhance control. Over time therefore, and across different intrinsic variables, reorganization may shift through the hierarchy, in a 'need-driven' way, to help reduce error where it is needed and help the system as a whole run smoothly. Indeed these are the basics of infant development according to PCT (Plooij, 2020). The focus of consciousness is not under volitional control according to this account, but attracted by whatever perceptual variable is in persistent error. Memory allows this process to also occur within spontaneous imagination. Yet, humans also appear to be able to sometimes *choose* to focus their conscious attention on certain aspects of their surroundings, and on specific tasks and skills, even when this is not needed to reduce intrinsic error. People follow routines, advice, and plans and they use logic and knowledge to decide what to focus on even when that experience does not capture their attention at that moment. This is where tertiary consciousness comes in.

## Tertiary Consciousness: Cooperation and Competition between Intrinsic and Propositional Systems for the Focus of Reorganization

In order to explain the extensive capacity of consciousness in humans, two further steps are required. The first is to explain how input signals can be maintained for long periods of time to allow reorganization to be most effective and to permit their 'categorization' within tertiary consciousness. The second follows on from the first to explain how humans bring aspects such as language, knowledge, plans, narrative, collective control, and logic to bear on conscious experience (see also Nevin, 2020).

An input signal is brief. To explain consciousness, there needs to be an explanation of how there is a phenomenal experience persisting over time. The simplest way to maintain an input is negative feedback control itself - varying action (e.g. body balance, eye movement) to keep focused on the same perceived aspect of the environment over time. Yet this is often still subject to perturbations, and information may need to be compiled over time to specify a perceptual variable (e.g. to identify an animal concealed in bushes). Several theories of consciousness propose that perceptual inputs need to be stored in a buffer memory and this memory continuously regenerated, if they are to enter and remain in focal consciousness (e.g. Barnard & Teasdale, 1991). One proposed method is to allow signals that enter at different times through different sensory pathways to be delayed with respect to one another ('delay lines') with smoothed temporal offsets, such that they are experienced coherently, within a 'buffer' zone (Barnard & Teasdale, 1991). Such a process is consistent with PCT, which incorporates delays, temporary integration, and memory retrieval as imagined perception into its architecture (p. 33, p. 208, Powers, 1973). This sustained input signal may then be subject to control by two distinctive sets of systems - intrinsic (covered earlier) and propositional.

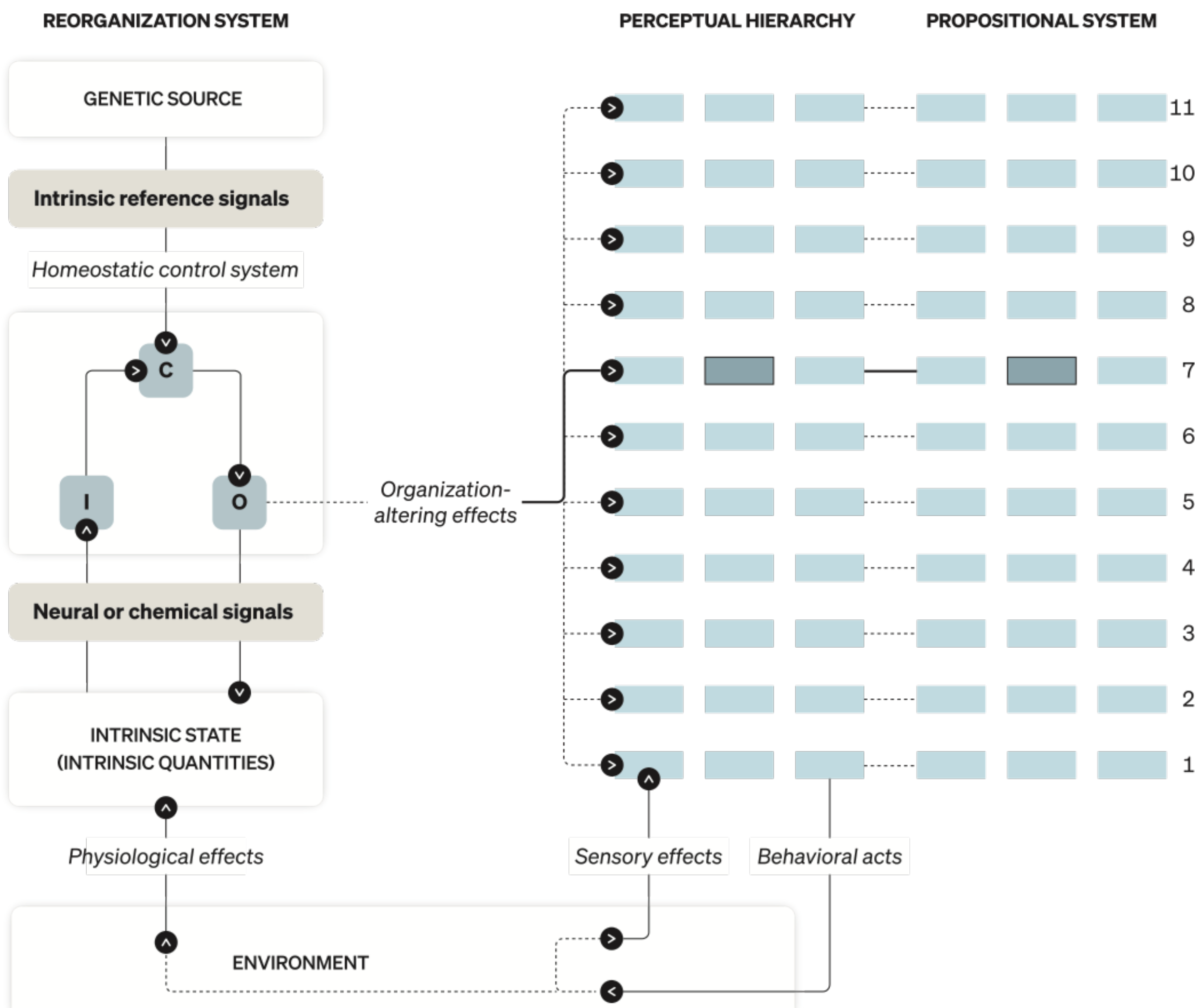
The specific profile of intrinsic control systems within an individual and their specific control parameters may specify the biological factors that to some degree influence their unique temperament and personality (Van Egeren, 2009). Furthermore, alterations in these biological systems through changes in the circadian cycle, drugs and medication will have an impact on their current intrinsic error. The implication here is that the intrinsic control systems within each individual are responsible for the various kinds of perceptions that are prioritized within consciousness on different occasions. For example, intrinsic systems governing nutritional state (e.g. glycemic control) will attract consciousness and possible reorganization towards food-related perceptions. The innate biological 'signature' or 'rhythm' will be unique to each individual and the perceptual hierarchy needs to develop in a manner over time that meets the needs of these multiple intrinsic systems (see also Leonard, 1981).

Conversely, humans have developed sophisticated systems of symbols that ‘represent’ specific perceptual variables. Symbolization is an ‘order reduction’ method (p. 318, Powers et al., 1960b) that has multiple benefits for individuals and for human culture as a whole. It allows elements to be combined and reorganized sequentially within language or as if-then plans in more elaborate ways than the perceptions themselves, and communicated to others at high fidelity to facilitate shared pursuits and verbal learning. Among the propositional systems in use are semantic knowledge, autobiographical knowledge, sciences, languages, cultures and religions. These systems include the ‘beliefs’ described by some predictive processing accounts. The relevance here is that sustaining a perceptual input in consciousness may be necessary to link it with a specific symbol. Once learned, this allows the individual to bring and sustain a perceptual signal to consciousness via its symbol (e.g. by writing down one’s thoughts), or indeed to remove it from consciousness by switching to another line of thinking. This enables humans to learn new roles, abilities, skills and tasks in systematic ways, often by focusing conscious awareness on each separate perceptual variable involved in the skill to allow improved control, and then shifting to another. This is typical of sports coaching, as is the introduction of a variety of obstacles and challenges to maximize learning, which in PCT terms is the enhancing of perceptual control despite disturbances. Within PCT, task analysis reveals the hierarchy of perceptual variables, each of which can be learned through being the focus of reorganization in consciousness (Marken, 1999). Indeed, a mobile focus of conscious awareness throughout a control hierarchy may be a marker of good psychological well-being (Watkins, 2011). The propositional system can thereby foster the formation and control of new perceptual variables - it can allow the planned engagement in creative pursuits. Thus, the propositional system can work in harmony, or against, the intrinsic systems governing the focus of consciousness at any one moment.

Yet, there are also potential disadvantages of propositional coding if it is not used in synergy with other control systems. First, logic and language can give the illusion of an agreed perception of reality between two or more individuals by ‘smoothing over’ the divide between them; one person's perception of a 'house', or 'justice', may be different from another person's, especially if they grow up in different cultures. Second, and relatedly, through attempting to control only symbols, the practical limitations of putting a plan into practice can be missed or poorly conceived - this is the classic divide between theory and implementation. Third, and importantly for the conscious individual, a plan, knowledge, or 'facts' that are mutually agreed within a propositional format may conflict with the intrinsic goals of the individual, for example within societal policies that are discriminative, or at an extreme, within harmful religious practices or suicide pacts.

Ultimately, all three systems need to work together for effective control (see Figure 7): only the perceptual hierarchy contains learned perceptual variables for control by bodily movement; only the intrinsic systems ensure that the individual's unique biological needs are met through reorganizing the perceptual hierarchy; and only the propositional system can direct conscious awareness to perceptual variables that are not currently entailing intrinsic error, and that may need to be reorganized in order to fit plans, achieve longer term goals, or maintain culturally appropriate values. Intrinsic systems control essential variables (e.g. osmolite concentration) and the intrinsic system experiencing unresolved error specifies the current interoceptive variable to control (e.g. the feeling of thirst) to guide reorganization at a higher level of the perceptual hierarchy (e.g. to learn locations of drinking water). Propositional systems form symbolic representations of perceptual variables for knowledge acquisition (e.g. bottles hold water), long-term plans (e.g. filling up a bottle for a long trip), communication (e.g. asking someone to pass a bottle) and recursive re-integration (e.g. identifying what a new water bottle looks like) within the hierarchy. Adaptive control in humans requires ongoing and mutual cooperation of both intrinsic and propositional systems within consciousness to reorganize the perceptual control hierarchy optimally. This *is* conscious human life - the adaptive switching between plans and multiple instincts, using conversation, language and the realm of cultural artefacts to establish and maintain control within the bounds of biological limits. Through working within society that develops over time, humans are able to utilise consciousness to broaden the range of variables in their environment they can control - humans strive to be a ‘universal perceptual control system’.

***Figure 7. A diagram to illustrate how logical/propositional systems are engaged in mutual control with intrinsic/homeostatic systems over the focus of reorganization within the perceptual hierarchy. The horizontal lines between the systems are a simplified way of showing the bidirectional, branching connections between them. In this example, both systems are maintaining reorganization, and conscious awareness, at the 'middle' unit on level 7 of the hierarchy. However, at other times they will conflict, such as when an intrusive perception is avoided in order to sustain concentration on a current task, For simplicity, this diagram also does not show the multiple intrinsic control systems, or the connections between units of the perceptual hierarchy (see Figure 2).*****Addressing**




## Challenges to Face Validity of the Model

I have provided a theoretical and empirical basis of the proposed model of consciousness. But does it have face validity? One would not expect conscious experience itself to transparently reveal the mechanisms of consciousness in the model, because if this were to be the case, no theories of consciousness would be required. Yet it is reasonable to expect that conscious experience itself would be consistent with the model and even provide some surface evidence of its mechanisms. Therefore, three challenges to the face validity of the model will be introduced and addressed below.

**"Consciousness is not Conflict"**

The model proposes that the purpose of consciousness is to resolve conflict , and yet consciousness does not feel like conflict most of the time. How can this critique be addressed?

Conflict occurs at multiple levels. Yet, only conflict between control systems at higher levels in the hierarchy - such as conflicts between the principles of honesty versus kindness, or between programs such as going to the cinema versus watching TV at home - would be clearly experienced as conflict. Conflicts between the control of low level variables such as intensities or sensations manifest themselves differently. Examples of experiences resulting from lower-level conflicts include the awareness of the input signals that occur when one shifts in one's seat to get comfortable, or when one's eyes dart around a visual scene to establish what one is looking at. Indeed, even if a person is engaged in a 'high-level', skilled activity, such as typing, the successful execution of this activity relies upon widely distributed lower level systems in the hierarchy, and conflict, and therefore error, may build up in any one of these systems. One example is the pain that is felt during injury when opposing muscles on a joint are misbalanced (Crosier, 2004); the pain would draw awareness and allow the person to learn a new way of completing the activity through reorganisation.

Whilst conflict can undermine control, it is not conflict itself that defines the focus of conscious awareness, and the conscious individual may well not identify the process of conflict itself. Rather, consciousness is drawn to the control unit that is in sustained error which has built up because its outputs are not sufficiently successful in counteracting disturbances (Powers et al., 1960a). This often happens because of conflict, but it is not conflict that constitutes consciousness itself. Indeed, the input functions to specify novel perceptual variables may be proactive in their capacity to reduce conflict. For example, receiving adulation and praise through learning to perform a complex, original and challenging skill for a public audience may, at least temporarily, reduce conflicts regarding whether or not to avoid meeting new people. The idea that consciousness is drawn to error is also consistent with the literature on inattention blindness - the failure to notice a fully-visible, but unexpected object because attention was engaged on another task. This field of research has challenged the common assumption that people passively and consciously perceive the visual scene around them. Rather, people tend to report perceiving aspects of their environment that are novel but task-relevant - and therefore involved in error correction - as opposed to aspects of a visual scene that are clearly distinct, moving and yet not related to their current goal (Horstmann & Ansorge, 2016).

Of course, humans do have the capacity to observe the felt sense of conflict as well, and to describe it in words, such as being undecided or ‘in two minds’, but this is a feature of tertiary consciousness, rather than being a necessary component of primary or secondary consciousness.

**The 'blank wall' hypothesis**

A second critique of the model can be termed the 'blank wall': it might feel self-evident that a person could sit in an empty room and look at a blank wall and still remain conscious. It does not feel like there is 'novel integration of information' occurring during most of our everyday conscious experience. Yet, the model proposes that primary consciousness is sustained by dynamic adjustments that maintain a specific rate of integration. Note that the model does not predict that the person would simply lose consciousness if they were to stare at a blank wall for a period of time. Rather, the model predicts that the perception of continually homogenous input such as 'a blank wall in an empty room' would drift away from consciousness, to be replaced in consciousness by inputs from other perceptual systems that would serve to maintain the integration rate. Below I review the evidence that is consistent with this prediction of the model.

At first I tried personal experimentation, which the reader is welcome to try out at home. I attempted the 'blank wall experiment'. I went up close to the wall so that it filled my visual field, but I immediately noticed that the wall was textured in an array of interestingly shaped bumps. So, I did my best to focus on just one of these. As I did so, within seconds, I began to notice other perceptions such as a window in my peripheral vision, birds singing outside, and my breathing. These were perceptions at the fringe of my awareness that then came to occupy the focus, as described in the model. Despite many attempts, I could actually only manage to focus for a few seconds on one, homogenous perception without another perception slipping in.

Is focusing conscious awareness really such a challenge? The challenges of sustaining a conscious focus on a simple display was widely studied after World War 2 following the lapses in vigilance to the enemy within military personnel (Mackworth, 1948). In these studies, the ability to notice a ‘double jump’ in the secondhand of a clock at the centre of the visual field starts to drop off after around 7 minutes. However, there appear to be no studies attempting to time the duration that a perception can stay continuously within focal awareness in the absence of a concurrent task. It is notable however that the ability to sustain attention is a skill that takes intensive training to improve, such as through meditation (MacLean et al., 2010). This is consistent with the proposal that an intrinsic system governing novel information integration naturally interferes with the capacity to maintain consciousness on one perceptual variable.

To check whether I was unusual in not being able to focus my conscious awareness for longer than a few seconds, I looked for more participants. Arguably, if we wish to understand primary consciousness, then children are more appropriate research participants, because they have not fully developed the voluntary control abilities that come with tertiary consciousness. So, I asked my sons, aged 6 and 10 years to do the blank wall experiment. They were much worse than me. They couldn’t manage it at all without talking, fidgeting, thinking about how much time they are wasting, or moving their eyes. Research indicates that fidgeting and mind wandering are closely correlated, maybe indicating they serve a similar function (Carriere, Seli, & Smilek, 2013). It might well be a truism that children get distracted easily and ‘can’t sit still’. This phenomenon seems mainly to be studied in the clinical context, as a sign of an ‘attentional deficit’. Yet it has been recognised as ubiquitous in young children. For example, it impacts on the accuracy of functional magnetic resonance imaging (fMRI). The inability of children to tolerate a monotonous environment without self-controlled stimulation (e.g. fidgeting & shifting position) in the scanner has spawned a range of interventions to make fMRI scans in children possible (Woods-Frohlich, Martin, & Malisza, 2010).

A related source of evidence in adults comes from studies of the effects of sensory deprivation. This involves an experimental apparatus that reduces sensory input as much as possible, such as by lying in a floatation tank within dark, soundproof room. A related approach is sensory homogenisation, which involves presenting continuous, monotonous or unstructured inputs, for example through goggles that present a constant red light and headphones that present white noise. The results of these experiments are stark and have been known for many years (Solomon, Leiderman, Mendelson, & Wexler, 1957). After several minutes, participants begin to report affective, somatosensory, vestibular, and bodily perceptions (Tsuji, Hayashibe, Hara, & Kato, 2004), and even vivid, dream-like mental imagery (Pütz, Braeunig, & Wackermann, 2006). These findings are consistent with the model: a control system is using shifts in attention and trial-and-error reorganisation to develop and modify the functions that integrate diverse sensory inputs in order to maintain an intrinsic integration rate.

## Brief Research Review and Proposals

The integrative control theory model presented here is clearly multi-componential and designed to lead to a working model as opposed to isolated hypotheses. Nonetheless, it does lend itself to empirical testing.

One hypothesis generated by the PCT model is that the functions and parameters of control systems controlling a specific perceptual variable will be subject to change when consciousness is directed towards them. An early study provided participants with a task in which they got higher scores when they could identify more complex, hierarchically superior, perceptual variables (Robertson & Glines, 1985). Whilst this study did not enquire about consciousness specifically, participants showed stepped improvements on the task as they learned to control new, hierarchically superordinate, perceptual variables in the task. Powers (1980), and in turn, Marken (2021) have proposed that the hypothesis could be tested more directly within a simple tracking task. The participant would be instructed to keep a cursor next to a fixed target on the screen, by moving a joystick to counteract the computer’s movements.

Time-varying pseudorandom disturbances would be applied to both the cursor movement and the resistance from the joystick throughout the task. A two-level computational model would be constructed that varies proprioceptive perception of force on the joystick at the lower level in order to control position of the cursor at the upper level. The computational model would be optimised to maximise fit with the experimental data at regular intervals as the participant is instructed to shift their focus of consciousness. A shift of the participant's focus of conscious towards the cursor should entail a better fit with models that randomly alter their parameters at the upper level, whereas a shift in focus of consciousness to the perception of effort in the hand should entail a better fit with models that randomly alter their parameters at the lower level.

A new hypothesis of this revised PCT account of consciousness is that novel information integration rate is a controlled variable within conscious individuals. A test for the controlled variable could be used, which is a robust scientific method to study purposive agents (Marken, 2021). If the environment is disturbed in a way that shifts the provision of information to potentially integrate novel information at a different rate, then the individual will act against this disturbance to keep the rate at the reference value, providing their intrinsic errors stay relatively constant (e.g. they are not in current danger). Examples of disturbances that lower the potential resource for novel information integration include environments with highly familiar, or very simple, environments. If novel information integration rate is a controlled variable then individuals should counteract the effect of these environments via one or more of the three systems. For example, given the opportunity, participants may shift their focus of visual attention towards complex but potentially integrated environments for which perceptual variables can be constructed and reorganized. A standardized attentional paradigm could be used to assess this shift (Fagan & Vasen, 1997). Alternatively, a computer paradigm could be developed that allows the individual to vary the informational content of their sensory input continuously (e.g. via the complexity of a dynamic visual image on the screen). Using this approach, the specific parameters for the control of information integration rate (e.g. reference value, gain) that are specific to an individual participant could be estimated and tested for their specificity. The testing of person-specific computational models has been applied successfully to manual tracking of a target (Parker, Tyson, Weightman, Abbott, & Mansell, 2017), but it would be new to consciousness research. Whilst there is evidence that some people vary on a dimension of ‘sensation-seeking’ (Roberti, 2004) or ‘stimulus-hunger’ (Suedfield, 1981), the studies establishing this evidence do not test for the control of novel information integration rate. Their findings are based on the correlation between psychophysiogical variables and self-reports of personality style within groups of participants, rather than model specification and testing of individuals within a task (for more general issues with these methodologies in the light of closed-loop control, see Mansell & Huddy, 2018).

In the absence of informational resource from the external environment, people may experience increased mind wandering as their propositional systems attempt to generate their own perceptual inputs for integration. Mind wandering is generally observed across a range of low information environments (Smallwood & Schooler, 2015) and may account for road accidents during familiar driving scenarios (Lin et al., 2016). If mind wandering is insufficient, people may experience sensory distortions and hallucinations, as their intrinsic systems attempt to restore information integration rate by enhancing trial-and-error reorganization. This latter finding in response to environments of sensory deprivation has been long established within experimental studies (e.g. Solomon, Leiderman, Mendelson, & Wexler, 1957).

Examples of disturbances that raise the potential for information integration above the reference level would be those that are novel and unfamiliar but highly complex and potentially meaningful, such as starting at a new school or college, going to a new country with a different culture, or watching a new movie or mini-series. Examples of processes to counteract these disturbances might include attempts to slow down, pause or avoid these situations or focus on simple, repetitive features of them. Thinking would slow down or become focused on simple, familiar themes. Intrinsic systems may restore information integration to a low enough rate by simplifying perceptual input, which might be observed as absent mindedness, overgeneralizing and missing the detailed features of a situation.

A similar hypothesis was put forward in the context of aesthetic experience (Van de Cruys, Bervoets, & Moors, 2021, p9): “It is easy to see our creative, artistic explorations as (one of) the system’s efforts to upregulate the rate to match the expected one. A concrete recent example was provided by the Covid-19 pandemic lockdown, during which people apparently en masse took up learning to play an instrument”). At the other end of the spectrum, some environments may raise uncertainty levels too quickly too much, which also prevents expected structuring rates from being re-established. Instead of exploration, here one would see (anxious) avoidance or repetitive, rigidly structured (stereotypical) forms of self-stimulation to reinstate the expected rates of uncertainty reduction (as also found in clinical disorders; e.g., Van de Cruys et al., 2021). As explained earlier, there are important differences between the predictive processing and PCT models, and the Van de Cruys et al. (2021) proposal was not focused on modelling consciousness itself. Nonetheless, the fact that they appear to generate a similar hypothesis, make it important to clarify further hypotheses from the PCT account.

In addition to the above plans, additional research would be needed to test the full model with respect to how intrinsic and propositional systems engage in mutual control of the focus of consciousness to resolve conflict. This process is consistent with evidence from the coding of individual sessions of psychological therapy (Grzegrzolka, & Mansell, 2019; Higginson & Mansell, 2018). In these studies, the language used by individual clients tends to shift from describing difficulties with control, through sustaining awareness on goal conflicts and felt somatic experience, to describing a higher-level goal regarding the self and the world. However this evidence is descriptive; testing a complete, computational model of these interacting systems against quantitative data would provide the most robust test (Mansell & Huddy, 2018).

## Summary

In the current article, a PCT perspective on consciousness is provided in answer to a number of limitations (and to integrate the strengths) of other contemporary theories of consciousness. The role of consciousness within the original formulation of PCT (p318-321, Powers et al., 1960b; p200-201, Powers, 1973) was first described. It operationalizes consciousness as indicative of the current focus of reorganization within a control hierarchy. This allows error that remains uncorrected through action in the environment, and that entails error in intrinsic (homeostatic) systems, to be reduced through sustained, random, trial-and-error changes in the properties of control systems. Classically, these changes need to be focused and targeted on the system superordinate to compartmentalized systems that conflict with one another.

In the current account, I have proposed that the control unit in focus equates with primary consciousness to the degree that it integrates novel information for the purpose of generating and reorganizing a novel input function to allow the control of a new perceptual variable, especially in situations in which multiple intrinsic control systems are in error and need to be controlled from a new, higher level perspective. In a further application of the theory to secondary consciousness, I have proposed that novel information integration rate is a controlled variable that allows consciousness to be sustained continuously. This provides benefits such as enhanced accuracy of control and conflict resolution, but it can also entail experiential avoidance. Control of novel information integration rate is managed by counteracting disturbances to the rate through processes such as varying the focus of attention, moving to a new environment, engaging in imagined generation of inputs, or altering the rate of reorganization. The current PCT model also provides a necessary role for buffered memory and propositional coding that allows tertiary consciousness via the symbolization of perceptual variables. This facilitates communication and collective control with other individuals, and also the capacity to control the focus of conscious awareness independently of the automatic attraction by prolonged intrinsic errors. It is proposed that many of the products of human consciousness, such as art, culture, language and civilization, and many of its limitations such as mental health problems and group conflict, emerge from this complete working system within which reorganization of the perceptual hierarchy is under mutual influence by intrinsic needs and propositional rules and plans.

The hypotheses generated by the PCT model have been described, including emerging evidence. The model currently lacks any description of its neural substrates (for a contemporary integrative review, see Northoff & Lamme, 2020), and its ramifications within interpersonal interactions and at a societal level are not elaborated here. Nonetheless, PCT provides the opportunity to do so, given its wide interdisciplinary applications (Mansell, 2020).

## References

Argentim, L. M., Rezende, W. C., Santos, P. E., & Aguiar, R. A. (2013, May). PID, LQR and LQR-PID on a quadcopter platform. In *2013 International Conference on Informatics, Electronics and Vision (ICIEV) (pp. 1-6)*. IEEE.

Baars, B. J., & Franklin, S. (2009). Consciousness is computational: The LIDA model of global workspace theory. *International Journal of Machine Consciousness*, *1*(01), 23-32.

Barnard, P. J., & Teasdale, J. (1991). Interacting cognitive subsystems: A systemic approach to cognitive-affective interaction and change. *Cognition and Emotion, 5, 1-39.*

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L., & Boothby, E. J. (2012). Automaticity in social-cognitive processes. *Trends in Cognitive Sciences, 16(12),* 593-605.

Barter, J. W., & Yin, H. H. (2021). Achieving natural behavior in a robot using neurally inspired hierarchical perceptual control. *Iscience, 24(9),* 102948.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18(2),* 227-247.

Carey, T. A. (2018). Consciousness as control and controlled perception. *Annals of Behavioral Science, 4*, 2:3. doi:10.21767/2471-7975.100034.

Carey, T. A., Mansell, W., Tai, S. J., & Turkington, D. (2014). Conflicted control systems: the neural architecture of trauma. *The Lancet Psychiatry, 1(4),* 316-318.

Carriere, J. S., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67(1)*, 19-31.

Cerullo, M. A. (2015). The problem with phi: a critique of integrated information theory. PLoS computational biology, 11(9), e1004286.

Chalmers, D. (2007). The hard problem of consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 225–235). Malden, MA: Blackwell Publishing.

Clark, K. B. (2018). Possible origins of consciousness in simple control over “involuntary” neuroimmunological action. *Consciousness and Cognition, 61,* 76-78.

Clune, J., Misevic, D., Ofria, C., Lenski, R. E., Elena, S. F., & Sanjuán, R. (2008). Natural selection fails to optimize mutation rates for long-term adaptation on rugged fitness landscapes. *PLoS Computational Biology, 4(9),* e1000187.

Corning, P. A. (2007). Control information theory: The ‘missing link’in the science of cybernetics. *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research, 24(3)*, 297-311.

Croisier, J. L. (2004). Muscular imbalance and acute lower extremity muscle injuries in sport. *International SportMed Journal, 5(3),* 169-176.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (1992). *Optimal experience: Psychological studies of flow in consciousness.* Cambridge University Press.

Cziko, G. (1997). *Without miracles: Universal selection theory and the second Darwinian revolution.* MIT press.

Cziko, G. (2000). *The Things We Do: Using the Lessons of Bernard and Darwin to Understand the What, How, and Why of Our Behavior.* MIT press.

Deane, G. (2021). Consciousness in active inference: Deep self-models, other minds, and the challenge of psychedelic-induced ego-dissolution. *Neuroscience of Consciousness, 2021(2),* niab024.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79(1-2),* 1-37.

Doerig, A., Schurger, A., Hess, K., & Herzog, M. H. (2019). The unfolding argument: Why IIT and other causal structure theories cannot explain consciousness. *Consciousness and Cognition, 72,* 49-59.

Fagan, J., & Vasen, J. H. (1997). Selective attention to novelty as a measure of information processing across the lifespan. In J. Burack & J. Enns (Eds.), *Attention, development, and psychopathology (pp. 55-73).* New York Guilford Press.

Forssell, D. (2016). *Perceptual Control Theory: An Overview of the Third Grand Theory in Psychology Introductions, Readings, and Resources*. http://www.livingcontrolsystems.com/download/pct_readings_ebook_2016.pdf

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". Cognition, 50(1-3), 115132.

Goekoop, R., & de Kleijn, R. (2021). How higher goals are constructed and collapse under stress: A hierarchical bayesian control systems perspective. *Neuroscience and Biobehavioral Reviews, 123,* 257–285. https://doi.org/10.1016/j.neubiorev.2020.12.021

Gray, J. A. (1995). The contents of consciousness: A neuropsychological conjecture. *Behavioral and Brain Sciences, 18(4),* 659-676.

Grzegrzolka, J., & Mansell, W. (2019). A test of the feasibility of a visualization method to show the depth and duration of awareness during Method of Levels therapy. *the Cognitive Behavior Therapist,* 12.

Hancock, P. A., & Weaver, J. L. (2005). On time distortion under stress. *Theoretical Issues in Ergonomics Science, 6(2),* 193-211.

Healey, A., Mansell, W., & Tai, S. (2017). An experimental test of the role of control in spider fear. *Journal of Anxiety Disorders, 49,* 12-20.

Higginson, S., & Mansell, W. (2018). The development and evaluation of the Depth and Duration of Awareness Coding Scheme (D-DACS). *The Cognitive Behavior Therapist, 11.*

Horstmann, G., & Ansorge, U. (2016). Surprise capture and inattentional blindness. *Cognition, 157,* 237-249.

Huether, G., Doering, S., Rüger, U., Rüther, E., & Schüssler, G. (1999). The stress reaction process and the adaptive modification and reorganization of neuronal networks. *Psychiatry Research, 87(1),* 83-95.

Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders, 36(1),* 77-90.

James, W. (1890). *Principles of psychology.* New York: Holt.

Johnson, T., Siteng, Z., Cheah, W., Mansell, W., Young, R., & Watson, S. (2020). Implementation of a perceptual controller for an inverted pendulum robot. Journal of Intelligent & Robotic Systems, 99(3-4), 683-692.

Kim, H., Hudetz, A. G., Lee, J., Mashour, G. A., Lee, U., Avidan, M. S., ... & ReCCognition Study Group. (2018). Estimating the integrated information measure phi from high-density electroencephalography during states of consciousness in humans. Frontiers in Human Neuroscience, 12, 42.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. Consciousness and Cognition, 9(2), 149-171.

Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making sense of the world: infant learning from a predictive processing perspective. *Perspectives on Psychological Science, 15(3),* 562-571.

Kotchoubey, B. (2018). Human consciousness: Where is it from and what is it for. Frontiers in psychology, 9, 567.

Layzer, D. (1975). The arrow of time. Scientific American, 233(6), 56-69.

Leonard G (1981) The Silent Pulse. Bantam Books, New York.

Lin, C. T., Chuang, C. H., Kerick, S., Mullen, T., Jung, T. P., Ko, L. W., ... & McDowell, K. (2016). Mind-wandering tends to occur under low perceptual demands during driving. *Scientific Reports, 6(1),* 1-11.

Lombardi, O., & López, C. (2018). What does ‘information' mean in Integrated Information Theory?. *Entropy, 20(12),* 894.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1(1),* 6-21.

MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., ... & Saron, C. D. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science, 21(6),* 829-839.

Mansell, W. (2020). Ten vital elements of perceptual control theory, tracing the pathway from implicit influence to scientific advance. In *The Interdisciplinary Handbook of Perceptual Control Theory* (pp. 585-629). Academic Press.

Mansell, W., & Carey, T. A. (2013). Dissociation: perceptual control theory as an integrative framework for clinical interventions. In *Cognitive behavioral approaches to the understanding and treatment of dissociation* (pp. 239-253). Routledge.

Mansell, W., Carey, T. A., & Tai, S. J. (2015). Classification of psychopathology and unifying theory the ingredients of a Darwinian paradigm shift in research methodology. *Psychopathology Review, 2*(1), 129-153.

Mansell, W., & Huddy, V. (2018). The assessment and modeling of perceptual control: A transformation in research methodology to address the replication crisis. *Review of General Psychology, 22(3),* 305-320.

Marken, R. S. (1999). PERCOLATe: perceptual control analysis of tasks. *International Journal of Human-Computer Studies,* 50(6), 481-487.

Marken, R. S. (2021). *The Study of Living Control Systems: A Guide to Doing Research on Purpose.* Cambridge University Press.

Markman, K. D., Klein, W. M., & Suhr, J. A. (2012). *Handbook of Imagination and Mental Simulation.* Psychology Press.

Marvan, T., & Havlík, M. (2020). Is Predictive Processing a Theory of Consciousness?. http://philsci-archive.pitt.edu/18426/

Matić, A., Valerjev, P., & Gomez-Marin, A. (2021). Hierarchical control of visually-guided movements in a 3D-printed robot arm. *Frontiers in Neurorobotics, 149.*

Mayr, U. (2004). Conflict, consciousness, and control. *Trends in Cognitive Sciences, 8(4),* 145-148.

McReynolds, P. (1971). The three faces of cognitive motivation. In Day, H. I., Berlyne, D. E., Hunt, D. E. (Ed.), *Intrinsic Motivation: A New Direction in Education* (pp. 33–45). Holt, Rinehart and Winston of Canada.

Montemayor, C., & Haladjian, H. H. (2015). Consciousness, attention, and conscious attention. MIT Press.

Nevin, B. (2020). Language and thought as control of perception. In *The Interdisciplinary Handbook of Perceptual Control Theory (pp. 351-459)*. Academic Press.

Newman, J., & Baars, B. J. (1993). A neural attentional model for access to consciousness: A global workspace perspective. *Concepts in Neuroscience*, *4*(2), 255-290.

Noble, D. (2008). Claude Bernard, the first systems biologist, and the future of physiology. *Experimental Physiology, 93(1),* 16-26.

Northoff, G., & Lamme, V. (2020). Neural signs and mechanisms of consciousness: is there a potential convergence of theories of consciousness in sight?. *Neuroscience & Biobehavioral Reviews, 118,* 568-587.

Oudiette, D., De Cock, V. C., Lavault, S., Leu, S., Vidailhet, M., & Arnulf, I. (2009). Nonviolent elaborate behaviors may also occur in REM sleep behavior disorder. *Neurology, 72(6),* 551-557.

Pütz, P., Braeunig, M., & Wackermann, J. (2006). EEG correlates of multimodal ganzfeld induced hallucinatory imagery. *International Journal of Psychophysiology, 61(2),* 167-178.

Parker, M. G., Willett, A. B., Tyson, S. F., Weightman, A. P., & Mansell, W. (2020). A systematic evaluation of the evidence for perceptual control theory in tracking studies. *Neuroscience & Biobehavioral Reviews, 112,* 616-633.

Parker, M. G., Tyson, S. F., Weightman, A. P., Abbott, B., Emsley, R., & Mansell, W. (2017). Perceptual control models of pursuit manual tracking demonstrate individual specificity and parameter consistency. *Attention, Perception, & Psychophysics, 79(8),* 2523-2537.

Pellis, S. M., & Pellis, V. C. (2017). What is play fighting and what is it good for?. *Learning & Behavior, 45(4),* 355-366.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioral control. *Progress in Neurobiology, 134,* 17-35.

Pikkarainen, E. (2021). Education, Consciousness and Negative Feedback: Towards the Renewal of Modern Philosophy of Education. *Philosophies, 6(2),* 25.

Plooij, F. X. (2020). The phylogeny, ontogeny, causation and function of regression periods explained by reorganizations of the hierarchy of perceptual control systems. In *The Interdisciplinary Handbook of Perceptual Control Theory* (pp. 199-225). Academic Press.

Powers, W. T. (1973). *Behavior: The control of perception.* Benchmark Press.

Powers, W. T. (1978). Quantitative analysis of purposive systems: Some spadework at the foundations of scientific psychology. *Psychological Review, 85(5),* 417-435.

Powers, W. T. (1980). A systems approach to consciousness. In *The Psychobiology of Consciousness* (pp. 217-242). Springer, Boston, MA.

Powers, W. T. (2008). *Living control systems III: The fact of control.* Bloomfield, NJ: Benchmark Publications.

Powers, W. T., Clark, R. K., & Farland, R. M. (1960a). A general feedback theory of human behavior: Part I. *Perceptual and Motor Skills, 11(1),* 71-88.

Powers, W. T., Clark, R. K., & McFarland, R. L. (1960b). A general feedback theory of human behavior: Part II. *Perceptual and Motor Skills, 11(3),* 309-323.

Rabinovich, Z., & Jennings, N. R. (2010). A hybrid controller based on the egocentric perceptual principle. *Robotics and Autonomous Systems, 58*(9), 1039-1048.

Roberti, J. W. (2004). A review of behavioral and biological correlates of sensation seeking. Journal of research in personality, 38(3), 256-279.

Robertson, R. J., & Glines, L. A. (1985). The phantom plateau returns. *Perceptual and Motor Skills, 61*(1), 55-64.

Rudrauf, D., Bennequin, D., Granic, I., Landini, G., Friston, K., & Williford, K. (2017). A mathematical model of embodied consciousness. *Journal of Theoretical Biology, 428,* 106-131.

Safron, A. (2020). An Integrated World Modeling Theory (IWMT) of consciousness: combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; toward solving the hard problem and characterizing agentic causation. *Frontiers in Artificial Intelligence, 3,* 30.

Schlicht, T., & Dolega, K. (2021). You can't always get what you want: Predictive processing and consciousness. *Philosophy and the Mind Sciences, 2.* philosophymindscience.org/index.php/phimisci/article/view/9153

Schweighofer, N., & Doya, K. (2003). Meta-learning in reinforcement learning. *Neural Networks, 16(1),* 5-9.

Selfridge, O.G. (1959) Pandemonium: a paradigm for learning. In Proceedings of the Symposium on Mechanisation of Thought Processes (Blake, D.V. and Uttley, A.M., eds), pp. 511–529, H.M. Stationary Office.

Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology, 2,* 395.

Shah, A. (2012). Psychological and neuroscientific connections with reinforcement learning. In *Reinforcement Learning* (pp. 507-537). Springer, Berlin, Heidelberg.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27(3),* 379-423.

Signorelli, C. M., Wang, Q., & Khan, I. (2021). A compositional model of consciousness based on consciousness-only. *Entropy, 23(3),* 308.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*, 487-518.

Solomon, P., Leiderman, P. H., Mendelson, J., & Wexler, D. (1957). Sensory deprivation: A review. *American Journal of Psychiatry, 114(4),* 357-363.

Suedfeld, P. (1981). Environmental restriction and “stimulus hunger”. In *Advances in intrinsic motivation and aesthetics (pp. 71-86).* Springer, Boston, MA.

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, *215*(3), 216-242.

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience, 17(7),* 450-461.

Tsuji, K., Hayashibe, K., Hara, M., & Kato, Y. (2004). Detailed analyses of Ganzfeld phenomena as perceptual events in stimulus-reductive situations. *Swiss Journal of Psychology, 63(3),* 217-223.

Tulving, E. (1985). Memory and consciousness. Canadian Psychology/Psychologie canadienne, 26(1), 1–12.

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review, 23(1),* 87-102.

Van de Cruys, S., Bervoets, J., & Moors, A. (2021). Preferences need inferences: Learning, valuation, and curiosity in aesthetic experience. https://psyarxiv.com/zh6nt/

Van Egeren, L. F. (2009). A cybernetic model of global personality traits. *Personality and Social Psychology Review, 13*(2), 92-108.

Vandekerckhove, M., Bulnes, L. C., & Panksepp, J. (2014). The emergence of primary anoetic consciousness in episodic memory. *Frontiers in Behavioral Neuroscience, 7,* 210.

Vandekerckhove, M., & Panksepp, J. (2009). The flow of anoetic to noetic and autonoetic consciousness: A vision of unknowing (anoetic) and knowing (noetic) consciousness in the remembrance of things past and imagined futures. *Consciousness and Cognition, 18(4),* 1018-1028.

Vannucci, L., Falotico, E., Di Lecce, N., Dario, P., & Laschi, C. (2015, July). Integrating feedback and predictive control in a bio-inspired model of visual pursuit implemented on a humanoid robot. In *Conference on Biomimetic and Biohybrid Systems (pp. 256-267).* Springer, Cham.

Watkins, E. (2011). Dysregulation in level of goal and action identification across psychological disorders. *Clinical Psychology Review, 31(2),* 260-278.

Woods-Frohlich, L., Martin, T., & L Malisza, K. (2010). Training children to reduce motion and increase success of MRI scanning. *Current Medical Imaging, 6(3),* 165-170.

Yaron, I., Melloni, L., Pitts, M., & Mudrik, L. (2021). How are theories of consciousness empirically tested? The Consciousness Theories Studies (ConTraSt) database. *Journal of Vision, 21(9),* 2195-2195.

Young, R. (2017). A general architecture for robotics systems: A perception-based approach to artificial life. *Artificial Life, 23(2),* 236-286.