Philosophical Psychology

ISSN: 0951-5089 (Print) 1465-394X (Online) Journal homepage: https://www.tandfonline.com/loi/cphp20

Beyond free will: The embodied emergence of conscious agency

Michael F. Mascolo & Eeva Kallio

To cite this article: Michael F. Mascolo & Eeva Kallio (2019) Beyond free will: The embodied emergence of conscious agency, Philosophical Psychology, 32:4, 437-462, DOI: 10.1080/09515089.2019.1587910

To link to this article: https://doi.org/10.1080/09515089.2019.1587910

PHILOSOPHICAL PSYCHOLOGY
2019, VOL. 32, NO. 4, 437–462
https://doi.org/10.1080/09515089.2019.1587910

Routledge
Taylor & Francis Group

ARTICLE

Beyond free will: The embodied emergence of conscious agency

Michael F. Mascolo[a] and Eeva Kallio[b]

[a]Psychology, Merrimack College, North Andover, MA, USA; [b]Education, University of Jyväskylä, Jyväskylä, Finland

ABSTRACT

Is it possible to reconcile the concept of conscious agency with the view that humans are biological creatures subject to material causality? The problem of conscious agency is complicated by the tendency to attribute autonomous powers of control to conscious processes. In this paper, we offer an embodied process model of conscious agency. We begin with the concept of embodied emergence – the idea that psychological processes are higher-order biological processes, albeit ones that exhibit emergent properties. Although consciousness, experience, and representation are emergent properties of higher-order biological organisms, the capacity for hierarchical regulation is a property of all living systems. Thus, while the capacity for consciousness transforms the process of hierarchical regulation, consciousness is not an autonomous center of control. Instead, consciousness functions as a system for coordinating novel representations of the most pressing demands placed on the organism at any given time. While it does not regulate action directly, consciousness orients and activates preconscious control systems that mediate the construction of genuinely novel action. Far from being an epiphenomenon, consciousness plays a central albeit non-autonomous role in psychological functioning.

ARTICLE HISTORY

Received 9 December 2016
Accepted 8 August 2018

KEYWORDS

Consciousness; free will; agency; emergence; embodiment

1. Introduction

People experience themselves as *conscious agents*. In everyday life, we experience ourselves as having the capacity to choose our actions freely and exert conscious control over them (Monroe & Malle, 2010). However, humans are also biological organisms composed of physical material. How is it possible for physical organisms to be conscious agents? Is it possible to reconcile the idea of conscious agency with the view that humans are biological creatures subject to material causality (Hateren, 2015; Mele, 2009; Pacherie, 2014)? In this paper, we propose an embodied model of conscious agency that explains how it is possible for humans to be conscious agents within a materialist worldview.

CONTACT Michael F. Mascolo michael_mascolo@yahoo.com Psychology, Merrimack College, 315 Turnpike Street, North Andover, MA 01845, USA

The task of developing a scientifically viable model of conscious agency is complicated by three perennial problems. The first relates to the legacy of mind-body dualism. Most philosophers have long abandoned the traditional Cartesian dichotomy between an incorporeal mind and a physical body. As monists, they reject the idea that psychological processes have any nonphysical basis. Nonetheless, psychological scholars routinely invoke a duality between "mental" and "physical" processes. In this regard, the meaning of the term "mental" remains vague. While philosophers and psychologists identify "mental" states in terms of phenomenal, cognitive, and related processes and qualities, there remain serious questions about how the "mental" emerges from the biophysical, and how it is possible for "mental" acts to cause "physical" behavior. One problem for a model of conscious agency is explaining what it means to speak of "mental" or psychological processes, and how they arise from and operate within biophysical systems.

Second, the concept of conscious agency implies that humans have some capacity for *free will*. At its most basic level, free will implies that individuals have the capacity to exert conscious *control* over their actions (Fischer, 1994; Mele, 2009) in the absence of external impediments (Hume, 1748/2008). Beyond this, philosophers have suggested that an action is *free* if it *originates from the individual* in the sense of being "up to the person" (Pereboom, 2001) or if a person has the capacity to *have acted otherwise* (Frankfurt, 1969). Still further, some hold that free will involves the capacity to produce genuine *novelty* – that is, acts not fully determined by past and present forces (Ekstrom, 2000; Kane, 1996). Thus, a second task for a model of conscious agency is to explain what it means to say that persons can control actions – freely or otherwise – under the presupposition of a materialist worldview.

A third issue concerns the role of consciousness in psychological functioning in general and in the control of action in particular. A large literature demonstrates that many activities traditionally believed to require consciousness can occur without conscious participation (Bargh, 1997; Hassin, 2013). Libet's (Libet, 1999; Libet, Gleason, Wright, & Pearl, 1983) well-known experiments showed that brain-based action potentials assumed to initiate simple motor responses were initiated before individuals became conscious of the urge to execute those responses. While controversial (Schlegel et al., 2013), these and related findings (Soon, Brass, Heinze, & Haynes, 2008) challenge prevailing ideas about the role of conscious intentions, free will, and "the person" in executing motor action. The third task for a model of conscious agency is to explain the role of conscious and nonconscious processes in the regulation of action.

In this paper, we develop a philosophical framework for understanding how psychological processes emerge as higher-order properties of complexly organized biological systems. Drawing on recent analyses of emergence

(Bedau, 2008; Sartenaer, 2016; Stephan, 2004), we offer the concept of *embodied emergence* to explain how psychological processes arise from biological processes without violating exclusion and causal-closure principles (Kim, 1998, 1999). Embodied emergence holds that psychological processes *are* higher-order biophysical processes that exhibit emergent properties – namely, the psychological capacities for *phenomenal experience* and the representation of *meaning*. Psychological properties are *irreducible* in the sense that they cannot be deduced from knowledge of their base elements. However, while irreducible, they are not *autonomous* in the sense of having causal powers that override those of their constituents. Because psychological processes *are* higher-order biological processes, the intractable "mind-body" problem reduces to a merely *difficult* "body-body" problem (Thompson, 2007) – namely, the question of how one set of biological process (i.e., those that mediate experience and the representation of meaning) play causal roles in relation to other biological processes (e.g., motor action).

Building upon this premise, we propose an embodied process model of psychological functioning which provides a framework for understanding the nature of consciousness and conscious agency. In contrast to everyday or lay conceptions of human agency (Monroe & Malle, 2010), consciousness cannot and does not function as an autonomous "center of control" (Hommel, 2007). While the capacities for experience and meaning are emergent properties of biophysical systems, the capacity for *behavioral regulation* is not. The capacity for self-regulation is an already existing capacity of living systems. Consciousness thus *transforms* but does not *create* the capacity for agency. It does so by allowing organismic systems to represent themselves and their worlds and thus regulates their actions on the basis of those higher-order representations. Consciousness functions to *integrate the outputs of preconsciously and affectively selected* representations of the most challenging demands presented to the organism at any given moment. As it arises and changes in moment-by-moment activity, consciousness continuously activates nonconscious constructive processes that mediate the assembly of novel strategies for meeting adaptive challenges. Consciousness thus *orient*s but does not directly control the construction of novel forms of action (Masicampo & Baumeister, 2011; Pacherie, 2014; Slors, 2015). Thus, far from being an epiphenomenon, consciousness plays a central, if non-autonomous, role in psychological functioning.

2. Lingering dualisms and the case for embodied emergence

One obstacle to developing a scientifically viable concept of conscious agency is the question of how psychological processes, including consciousness, arise from and play causal roles within biophysical systems. Figure 1 identifies a variety of positions related to the mind-body distinction. As shown in Figure 1, in attributing free will to a disembodied mind, Descartes crystalized the mind-body duality – a distinction with which we continue to struggle today. The problems of Cartesian dualism are well known. How can *mind* – viewed as an incorporeal thinking substance – affect or be affected by the *body* – something that, by definition, can only be influenced by material causes? In what follows, we examine current issues related to the concept of emergence. We offer *embodied emergence* as a framework for understanding how phenomenal properties arise and operate in biophysical systems without violating the principles of physical causality.

2.1. Emergence and reductionism

Reductionist approaches are based on a direct rejection of dualistic thinking. As shown in Figure 1, they maintain that psychological processes are physical processes that can be explained exclusively with reference to the properties and interactions of their components (Stephan, 2004). *Identity-based* reductionism holds that behavior and psychological states are identical to brain states (Smart, 1959). *Eliminativism* holds that everyday folk conceptions of mind are erroneous and lead us to seek mentalistic explanations for what are essentially brain processes (Churchland, 1986). *Epiphenomenalism* maintains that, since any mental state is already caused by a brain state, psychological processes can have no additional causal power of their own (Pockett, 2004). The benefit of reductionism is that it fits the traditional scientific conceptions of causal determinism. However,

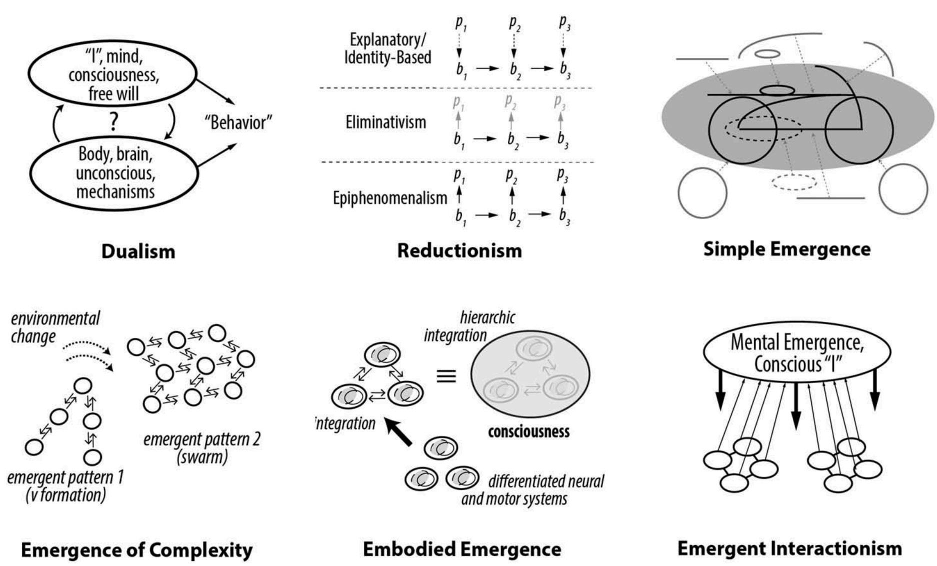

Figure 1. Reductionist and emergent alternatives to dualism.

it fails to account for our experience as *psychological* and *agentive* beings (Chalmers, 2006). Under reductionism, conscious agency is, at best, an illusion (Churchland, 1986; Wegner, 2002).

The concept of *emergence* provides an attractive way to resolve the mental-physical paradox (Anderson, 1972; Bedau, 2008; Chalmers, 2006). Emergence is based on the idea that collectives can exhibit *systemic properties* (Stephan, 2004). A systemic property is one that is exhibited by the system as a whole but is not found in its base elements. A common (albeit contestable) example involves how the liquid properties of water arise from the coupling of two gases – hydrogen and oxygen. Extending this simple idea to complex phenomena, consciousness is an emergent product of complexly organized neurobiological systems. Consciousness and related psychological properties arise from the integration of lower-order neurobiological systems that, by themselves, do not exhibit properties of conscious life.

If systemic properties were the only condition for identifying emergent phenomena, many phenomena would qualify as emergent. Theorists have proposed a series of conditions that are relevant to clarify the nature of emergent phenomena. These include the *irreducibility, novelty, unpredictability*, and *autonomy* of systemic properties (Bedau, 2008; Stephan, 2004; Sartenaer, 2016). The concept of novelty implies that emergent entities are not predictable, even in principle, from their antecedent conditions. Emergent phenomena are novel in the sense that they *come into existence* over time and do not exist prior to their creation. In this way, an emergent phenomenon is something that, at least before it arises, cannot be *predicted* from knowledge about its elements.

Perhaps the most important condition relevant to the concept of emergence is irreducibility. In this paper, we adopt Stephan's (2004) definition: a systemic property is irreducible if it is (a) *behaviorally unanalyzable* or if (b) the behavior of a system's components, as they function within the system itself, *does not follow* from the components' behavior as they function in isolation or in other less complex systems. An unanalyzable property is one that cannot be decomposed into further elements. Thus, although "washing dishes" is analyzable, perceiving a "red hue" is not. The second condition holds that a property is irreducible if its components function differently within a given collective or behave in a way that cannot be deduced, even in principle, from their behavior in less complex systems. Defined in this way, emergent properties have been studied both experimentally and in mathematical modeling (Boogerd, Bruggeman, Richardson, Stephan, & Westerhoff, 2005; Hoel, Albantakis, Marshall, & Tononi, 2016).[1]

2.2. Forms of emergence

Given these distinctions, it is possible to discriminate alternative versions of *emergence*. As shown in Figure 1, in *simple* or *nominal emergence* (Bedau, 2008), elements combine to produce a whole with systemic properties. However, the elements themselves retain their properties when they operate in isolation or as part of less complex systems. This form of emergence applies to simple mechanical devices. While the capacity to operate as a mode of transportation is an emergent property of the bicycle as a whole, this property is not true of its parts. Nonetheless, the parts of the bicycle retain their functional properties in isolation or as parts of less complex wholes.

Strong models of emergence maintain that systemic properties are novel and irreducible to those of their components (Boogerd et al., 2005; Hoel et al., 2016; Stephan, 2004). Some versions of strong emergence hold that as lower-order elements combine in complex ways, new wholes arise with properties that are autonomous in relation to their parts (see Wilson, 2015). As a result, they are able to exert downward causality over their base elements (Prosser, 2012; Sperry, 1991). To the extent that novel wholes are autonomous, their behavior *cannot be predicted* from knowledge of their components.

Sperry's (1976, 1980, 1991) *emergent interactionism* (Figure 1) provides an example of a model of strong emergence. Over the course of a long and distinguished career, Sperry argued for a conception of emergent mentalism without dualism. From this perspective, an irreducible mental level of functioning emerges from complex biophysical organizations. Once it emerges, it exerts downward control over its base elements. This view offered "renewed recognition of the role of mental *over* material forces" (Sperry, 1976, p. 11, emphasis added). Sperry was adamant that his view did not attribute supernatural, spiritual, or mystical properties to mental functioning. Nonetheless, through its capacity to stabilize and organize brain functioning, the causal powers of consciousness are stronger than those of its base elements (Sperry, 1991).

Many arguments have been leveled against this strong conception of a causally emergent consciousness. The *closed causality* argument holds that, to the extent that physical events have physical causes, physical causality must operate as a closed system: no cause that is not physical can exert an effect within a physical system. This rules out the possibility that non-physical mental events can affect physical processes. Similarly, the *causal exclusion* argument maintains that if consciousness – as an emergent "mental" process – arises from physical causes, it can have no causal import beyond the effects of physical causes that produce it. That is, any novel mental effects of conscious states can be attributed to the initial *physical* origins of those conscious states (Kallestrup, 2006; Kim, 1998, 1999). Theorists have proposed ways to address these problems. Models embracing *weak emergence* (Bedau, 2008) relax the requirement of strong irreducibility. From this point of view, systemic properties are fully dependent upon the arrangement of their component parts. The seeming irreducibility of systemic properties occurs as a result of the complexity of the dynamic interactions that occur among base elements. System elements function as dynamic systems (Kelso, 1995) whose elements *mutually influence* each other in nonlinear ways over time. In so doing, they *self-organize* into novel and unpredictable patterns involving both order and chaos.

Models based on weak emergence explain the production of complexity through the "bottom-up" interplay of component processes. Such models are thus able to avoid the problems associated with strong conceptions of downward causation. Complex systems models are most suited for explaining emergent patterns at a *horizontal* level of functioning (e.g., flight patterns in birds, weather patterns, ocean currents and waves, supply and demand in economic systems). With exceptions (Kelso, 1995; Thompson, 2007), dynamic systems approaches have been less effective in explaining vertical emergence – that is, how phenomenal processes arise within biophysical systems. Indeed, to the extent that such approaches relax the notion of irreducibility, weak models of emergence are compatible with many versions of reductionism (Stephan, 2004).

2.3. Embodied emergence

We invoke the concept of *embodied emergence* to understand how consciousness and other psychological properties can emerge from biophysical processes and simultaneously play causal roles in the production of action. Consistent with *weak* versions of emergence, embodied emergence maintains that emergent properties are identical to the *integrations* that take place when lower-level component processes are coordinated into higher-order organizations (Campbell & Bickhard, 2011). This process is indicated in the bottom-center panel of Figure 1. In embodied emergence, the *hierarchical integration* of lower-order processes is the *equivalent* (≡) of the higher-order structure. Hierarchical integration refers to the process by which two or more lower-order structures are inter-coordinated into a single higher-level organization. This kind of integration is frequently invoked in models of biological and psychological development, where the development of anatomical or behavioral structures occurs through the successive differentiation of lower-order elements and their coordination into higher-order organizations (Kallio, 2011; Mascolo & Fischer, 2015; Piaget, 1985; Werner & Kaplan, 1963/1984).[2] That is, there are not (a) lower-order processes, (b) their integration, *and then also* (c) another something that emerges above and beyond that structural integration. Thus, psychological properties are not "mental" processes with properties and powers that "hover above" their biological constituents. The coordination of lower-order biological subsystems into higher-order biological structures creates a new biological organization – a new biological whole. This higher-order organization is the new biological whole, and it is that whole which exhibits emergent properties.

Thus, consistent with *weak* emergence, embodied emergence embraces the idea that systemic properties result from and thus rely upon the organization of base elements. However, consistent with models of *strong* emergence, we propose that novel wholes are nonetheless *irreducible* to their component parts in the sense articulated by Stephan (2004). That is, some higher-order organizations produce emergent properties that either (a) cannot be *behaviorally analyzed* into smaller elements or (b) have base elements that behave in novel ways when they operate as part of the higher-order system. With this proposition, however, we further discriminate between the irreducibility and purported autonomy (Wilson, 2015) of emergent forms. A structure can be irreducible to its component parts without being understood as having novel autonomous properties that can override those of its parts. We suggest that the assumption that psychological processes are *both* irreducible *and* autonomous makes the problem of downward causation an intractable one. The idea that an irreducible "mental" something emerges with autonomous powers that exert control *over* physical processes is essentially a form of mind-body dualism. Under such assumptions, we find ourselves attributing mysterious powers of "control" to an autonomous "mental" process.

In holding that emergent psychological processes are irreducible but not autonomous, embodied emergence maintains the following:

(1) Psychological processes emerge as higher-order forms of complexly organized biological processes. Psychological processes *are* higher-order biological processes, albeit with novel qualities.
(2) A psychological process is any organismic process that mediates or is mediated by phenomenal experience and meaning, where meaning consists of the schematization of phenomenal experience. In making these assertions, we avoid characterizing psychological processes as "mental" – a term that can suggest something other than a biophysical process. What makes a biological process also a psychological process is its role in the constitution and mediation of experience and meaning.
(3) Although phenomenal experience and meaning are emergent, the capacity for regulation or control, in itself, is not an emergent psychological property. Biological regulation is an already existing property of biological systems.
(4) As emergent processes, psychological processes allow organisms to operate on the basis of the *meaning* and *significance* of their circumstances, rather than on the basis of mere physical stimulation. We

suggest that consciousness, as intentional awareness, functions to *integrate* novel sources of information and plays an *orienting* rather than directing role in production adaptive action.

(5) Psychological processes, including consciousness, thus transform existing capacities for biological regulation; they do not *introduce* the capacity for control into the system. As higher-order biological processes with coordinating and orienting functions, they are able to play a causal (but not autonomous) role in the production of action.

To the extent that hierarchical regulation already exists in biological organisms, there is no need to postulate consciousness in order to explain the origins of control over behavior. This idea – coupled with the idea that psychological systems are themselves biological systems – frees us from the need to postulate emergent forms of "mental" control that violate principles of biophysical causation.

3. The biological roots of agentive control

Agency need not be seen as a mysterious process. At its most basic level, agency is a form of *self-regulation*. Self-regulation is a basic property of all living organisms. In biology, regulation and self-regulation are routinely invoked analyses of *homeostasis, set points, feedback loops, activation thresholds*, and so forth. Building upon such concepts, it becomes possible to speak of the capacity for control without concepts like *free will*, which implies the autonomy of consciousness in the control of action.

Bich, Mossio, Ruiz-Mirazo, and Moreno (2016) have proposed a series of principles for defining the concept of regulation in biology. Regulation occurs when (1) a regulated process (*C*) can be dynamically decoupled from a regulating (*R*) process, (2) regulating processes are instantiated in circumstances involving *perturbation* [*P*] in internal or external conditions, (3) regulation activates a shift in dedicated processes that perform particular organismic functions which, thereupon, (4) function in ways to cope with novel perturbations. The process of regulation involves the operation of feedback and feed-forward loops that modulate the functioning of system components and mediate relations between system elements and the system's functioning as a whole (Bich & Moreno, 2016; Bechtel & Abrahamsen, 2010).

3.1. Simple biological regulation

These principles can be defined and illustrated through an examination of the process of chemotaxis in single-celled bacteria. *E. coli* is a single-celled bacterium found in the intestines of endothermic (warm-blooded) organisms. Chemotaxis is the process by which the bacterium shifts movement toward or away from chemical gradients of its environment. The *E. coli* bacterium is composed of three major parts: (a) a series of receptors that detect chemo-attractors (e.g., amino acids, oxygen, glucose) and chemorepellents (inorganic salts, amino acids); (b) a flagellar system, composed of a set of flagella and the flagellar motor; and (c) a series of chemical pathways that bridge the two.

The receptor and chemical system and the flagellar system work as dynamically decoupled systems. This means that while each system is *capable* of working on its own, when coupled, one system (the receptor complex) regulates the other (the flagellar system). The regulating system is activated by changes in environmental conditions, namely, changes in the surrounding chemical gradient. In the absence of chemo-attractors, *run-and-tumble* movement occurs. A *run* consists of a long, straight movement caused by clockwise flagellar movement. Counter-clockwise flagellar movement produces a *tumble* – an abrupt and random rotation of the organism – which changes the direction of the organism’s movement, perhaps in the direction of a chemo-attractor gradient.

This process illustrates principles of regulation as described by Bich et al. (2016). In *E. coli*, the receptor and chemical system regulates (R) the movement of the flagellum (C) in terms of two dedicated operations – counter-clockwise (C') and clockwise (C'') movement. Changes in chemical gradients in the environment create perturbations, which stimulate the R process. At the level of the organism, these operations create a negative feedback loop: the absence of chemo-attractors results in random shifts in movement until the organism enters an environment containing chemo-attractors. Feedback from chemo-attractors inhibits run-and-tumble behavior and returns the organism to its default run behavior, which results in further movement toward chemo-attractors.

3.2. Hierarchical regulation in complex biological systems

In more complex organisms, self-regulation occurs through the operation of hierarchically organized control systems that operate according to principles of positive and negative feedback. Negative feedback occurs when regulatory processes (R) produce controlled changes (C) that inhibit the regulatory R process over time ($R \leftrightarrows C$); positive feedback occurs when regulatory processes (R) produce controlled changes (C) that further activate the regulatory R process over time ($R \leftrightarrows C$). In organisms with nervous systems, neural systems can function as higher-order regulating systems that regulate lower-order systems in terms of negative feedback. Circadian systems provide a paradigmatic example of a hierarchically nested, self-regulating system (Bechtel & Abrahamsen, 2010). Circadian systems govern rhythmic changes in biological and psychological systems over the course of a 24-h period. Circadian processes exist in virtually all cells of the body, and they are essential in the regulation of sleep and alertness, body temperature, immune responses, hormonal release, and many other processes over a 24-h period. Within individual cells, the core circadian process consists of a 24-h feedback loop (Herzog, 2007).

The circadian system provides a rich example of the process of hierarchical regulation of biological and psychological functions. In mammals, the *suprachiasmatic nucleus* (SCN) functions as the master higher-order regulator of the circadian networks of the body. The SCN contains approximately 20,000 nerve cells – each of which is governed by the circadian process described above – and is located in the hypothalamus. When individual SCN cells are experimentally isolated, they operate according to their individual circadian cycles. However, as part of the SCN, their rhythms become coordinated. The SCN has regulatory connections to multiple areas of the brain which themselves regulate the timing of still-lower-level processes over a 24-hour period.

Further, the hierarchical regulation of circadian systems extends beyond the individual organism. Circadian systems evolved within the context of light-dark cycles caused by the earth's 24-h rotation on its axis as it orbits around the sun. In this way, circadian systems are part of a larger ecological system. Within this larger system, circadian cycles are initially decoupled, in the sense that they are distinct systems that can function independently from one other. Within individuals, even though circadian rhythms operate on a 24-h clock, they must be entrained by the light-dark cycles of the local environment. Thus, dark-light cycles and the modulation of light through the retina, as well as the activity of the SCN, act as nested regulators that modify constraints on lower-level systems, even as each system operates according to its own endogenously produced cycles. In this way, circadian cycles illustrate how the dynamic coupling of distinct biological systems mediates the process of hierarchical regulation.

4. An embodied process model of conscious agency

Building on the preceding sections, we now outline an embodied process model of psychological functioning that provides a framework for understanding conscious agency (see Figure 2). We begin with the assumption that, as properties of higher-order biological processes, psychological processes mediate adaptive action by allowing organisms to represent the meaning of various aspects of their worlds (Thompson, 2007). We elaborate three propositions that will be developed throughout the paper. First, as a product of embodied emergence, conscious activity is necessarily a product of *nonconscious* biological processes. Second, while consciousness does not function as

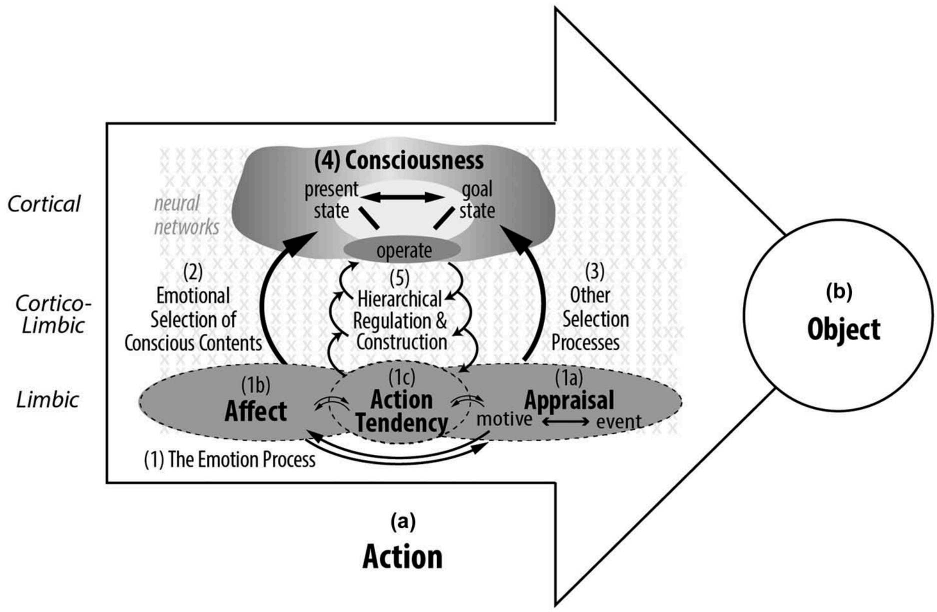

Figure 2. The architecture of conscious agency.

an autonomous controller of action, it integrates novel aspects of pressing adaptive challenges. Third, as adaptive challenges become represented in awareness, consciousness continuously activates preconscious constructive processes that bootstrap novel solutions to adaptive challenges. Thus, consciousness plays a central but non-autonomous role in integrating, orienting, and activating meaning and action.

In what follows, we define consciousness as the *structure of awareness*. Consciousness is *intentional* in the sense that it is *about* or directed *toward* something (Brentano, 1874/1995). We are not simply conscious; we must be conscious *of* something. The intentionality of conscious action is indicated in Figure 2 by the large arrow (Point A) and its object (Point B). We further differentiate between focal and peripheral consciousness. At any given moment, focal consciousness (attention) consists of that of which a person is *directly aware of* or that to which the person is *consciously attending*. This is indicated in Figure 2 by the white bubble within Point 4 (consciousness). Peripheral consciousness consists of the unattended surroundings of focal awareness (Lamme, 2003). We use the term *experience* as an umbrella term that refers to the full range of phenomenal qualities available to consciousness.

4.1. Emotion and the preconscious construction of conscious awareness

In everyday life, we tend to experience our intentions as springing spontaneously from within consciousness. This likely contributes to our experience of ourselves as capable of exerting conscious control over action "free from internal or external constraints" (Monroe & Malle, 2010, p. 215). However, conscious intentions must have causal origins. Unless consciousness functions as its own uncaused cause, intentions must have their origins outside of consciousness. Many forms of preconscious selection participate in the organization of consciousness (Points 2 and 3 in Figure 2). These include selective attention (Lamy, Leber, & Egeth, 2013), unconscious priming (Agafonov & Shabad, 2010), and unconscious effects of motor action (Halász & Cunnington, 2012). *Emotion* (Point 1 in Figure 2) plays a particularly important role in organizing consciousness.

In any given situation, amidst the enormous complexity of perceptual experience, only certain aspects of our worlds become conscious. For example, while operating an automobile we often find ourselves unaware of aspects of our driving (Charlton & Starkey, 2011). However, we become *fully aware* when a child crosses our path. How is this possible? We must look for the answer outside of consciousness itself, as *awareness* cannot make itself aware of that of which it is *unaware*. A large body of research suggests that emotion plays an important role in organizing consciousness (Baumeister, Vohs, DeWall, & Zhang, 2007; Tomkins, 1984). Contemporary research suggests that emotion is organized around three classes of mutually regulating systems (Point 1 in Figure 2). In any given context (see Point 1a), motive-relevant *appraisals* continuously and preconsciously monitor relations between circumstances and a person's motives and concerns (Moors, Ellsworth, Scherer, & Frijda, 2013; Roseman, 1991; Ruys, Stapel, & Aarts, 2011; Scherer, 2004). As preconscious appraisals encounter events relevant to a person's motives, they activate (see Point 1b) *affective* changes (i.e., feeling tone) and *bodily* changes (i.e., patterned physiological changes).

Emotional and preconscious appraisal processes mutually influence each other in the course of ongoing action (Lewis, 1996; Freeman, 2000). While preconscious appraisals generate affective changes (i.e., feeling tone, bodily transformations, and motive-related action tendencies), the resulting *affect* thereupon *selects* the results of pre-conscious appraisal activity for conscious awareness. Thus, pre-conscious appraisal activates affect, which thereupon *selects* and *organizes* these very same appraisals for conscious awareness (Lewis, 1996; Jarymowicz, 2010; points 2 and 4 in Figure 2). Feeling tone *amplifies* the significance of the appraised event (Tomkins, 1984) and marks it for learning and memory (Fiacconi, Dekraker, & Köhler, 2015). Affective processes not only modify feeling tone (Roseman, Wiest, & Swartz, 1994); they also activate and modulate (see Point 1c) *action tendencies* (Baumeister et al., 2007; Fontaine & Scherer, 2013), which function in the immediate service of adaptive challenges. For example, as pre-conscious appraisals register the impending danger of a child in the path of one's car, action tendencies are immediately evoked that function to remove the danger (e.g., breaking, swerving the car). Ongoing appraisal continuously modulates conscious awareness and affect-laden action tendencies until the adaptive challenge is ultimately resolved. In this way, appraisal, affect, and action *mutually regulate* each other (Lewis, 1996; Mascolo, Fischer, & Li, 2003) in real time. Thus, emotion plays a basic role in organizing conscious experience and action – intentional or otherwise (Freeman, 2000).

4.2. The integrative functions of consciousness

To say that consciousness arises from a pre-conscious constructive activity does not imply that consciousness is inert. A general agreement has emerged among consciousness researchers that Morsella (2005) has called "the integration consensus." This idea, common to many approaches (Baars & Franklin, 2007; Bonn, 2013; Edelman, 1989; Freeman, 2000; Morsella, Godwin, Jantz, Krieger, & Gazzaley, 2017; Tononi, 2013), holds that consciousness functions to integrate outputs of lower-order psychological systems that otherwise function nonconsciously, or more or less independently. Similarly, Baars and Franklin (2007) suggest that consciousness functions as a *global workspace*, bringing together different types of information from across the brain.

Research supports the idea that consciousness serves the functions of synchronizing and coordinating distributed systems of neural activity (Bola & Sabel, 2015; Rissman & Wagner, 2012; Tononi, 2007). However, what types of integrations are represented in consciousness? Research suggests that consciousness is not required for many forms of integrative activity (Mudrik, Breska, Lamy, & Deouell, 2011). As a result, Mudrik, Deouell, and Lamy (2012) maintain that it is not integration *per se* but the representation of *novelty* that is the function of consciousness. Consciousness is thus invoked when novel demands cannot be readily resolved by lower-level automatic and unconscious processes (Mudrik et al., 2012). Morsella, Godwin, Jantz, Krieger, & Gazzaley (2015) have suggested that conscious integration functions in the service of the skeletal-motor output system. From this view, consciousness mediates the connection between perceptual (and other) processes and motoric regulation. As discussed above, on occasions of adaptive challenge, affective processes organize representations of motivationally significant events in consciousness. Consciousness thereupon functions to bring together and coordinate "preconsciously vetted," motive-relevant meanings, experiences, and demands. Demands represented in consciousness are those that cannot be immediately resolved by existing endogenous resources and thus pose novel challenges. *Consciousness thus functions as a system for coordinating representations of the most pressing demands placed on the organism at any given time.*

4.3. Conscious integrations orient the construction of action

As they register and reverberate in consciousness (Grossberg, 2013), novel and goal-related contents of consciousness become increasingly organized, integrated, and consolidated. As a result, the goal-related contents of consciousness function as regulatory standards (R) that *orient* and *activate* lower-order control structures (C) (Carver, Johnson, Joorman, & Scheier, 2015; Powers, 2005). Control structures are hierarchically organized ($R \rightarrow C \rightarrow C' \rightarrow C'' \rightarrow C^x$). The hierarchical regulation of action has been verified in a variety of neurological and psychological domains, including the execution of basic motor skills (Gorniak, Zatsiorsky, & Latash, 2009), skilled actions such as driving (Medeiros-Ward, Cooper, & Strayer, 2014), and typewriting (Yamaguchi, Crump, & Logan, 2013). As higher-level adaptive demands are represented in consciousness, they function as top-level goals that orient the construction and operation of lower-level control systems (R → C). For example, at any given moment, higher-order conscious goals (e.g., *be a good student*) activate increasingly lower-order operations (e.g., *study for the test* → *re-read the book* → *turn the page*) that ultimately recruit motor control systems (e.g., *move hand* → *close fingers around page* → etc.). Perceptual feedback from motor action is continuously compared to higher-order reference standards (e.g., *changes in world* ⇆ *control fingers* ⇆ *move hand* ⇆ *turn page* ⇆ *read words* ⇆ *study*). This ultimately produces feedback to the highest level of representational and conscious activity. As long as higher-order goals are left unmet or unmodified, the regulatory process continues. When highest-level goals are met, they terminate or inhibit (—⊣) lower-order operations (e.g., *I am being a good student* ⇆ *study for test*), and the system moves onto other acts. The orienting role of consciousness in hierarchical regulation is indicated in Figure 2 at Point 5.

To the extent that consciousness activates lower-level control systems, it plays an important but indirect role in the regulation of action (Masicampo & Baumeister, 2013; Pacherie, 2014). The coordinating and organizing functions of consciousness can be understood as a form of "self-programming" (Slors, 2015). Conscious contents *orient* the operation of lower-order control systems, but they do not directly "trigger" motor action. Smooth execution of action occurs only when operations are already available to meet adaptive challenges. Novel adaptive challenges require the construction of *new* adaptive strategies. In the context of adaptive conflict – perturbations between goals and experience – the psychological system *equilibrates* (Grossberg, 2013): it activates nonconscious processes which assemble novel possible strategies for meeting ongoing challenges (Bola & Sabel, 2015; Di Paolo, Barandiaran, Beaton, & Buhrmann, 2014; Piaget, 1985).

Figure 3 shows a simple model of how an interplay between conscious and nonconscious processes can produce *genuinely novel* forms of thinking and acting. Imagine a thirsty individual who wants to keep her iced coffee cool without having it be diluted by melting ice. As this adaptive challenge is organized in consciousness (e.g., "I wish I could cool my coffee without ice!"), it activates distributed neural networks which underlie the construction of concepts such as *coffee, ice, melting*, and so forth. Activation spreads throughout networks that underlie the representation of related concepts. Consequently, new meanings related to the problem at hand become available to consciousness. For example, the thoughts "freezing makes water turn to ice" and "coffee is made from water" become *jointly represented* in consciousness. At this point, the *pair* of representations activates further memory processing, making multiple problem-relevant meanings available to consciousness (e.g., "coffee is made from water," "coffee freezes"). As new combinations of goal-related ideas reverberate in consciousness, *genuinely novel* representations and problem-solving strategies are consolidated – such as making ice cubes out of *coffee* rather than *water*.

5. How self-consciousness transforms human agency

Human conscious agency is transformed by the capacity for self-awareness. Research suggests a distinction between a *primary* pre-linguistic level of self-awareness and a *secondary* symbolically mediated and reflective level

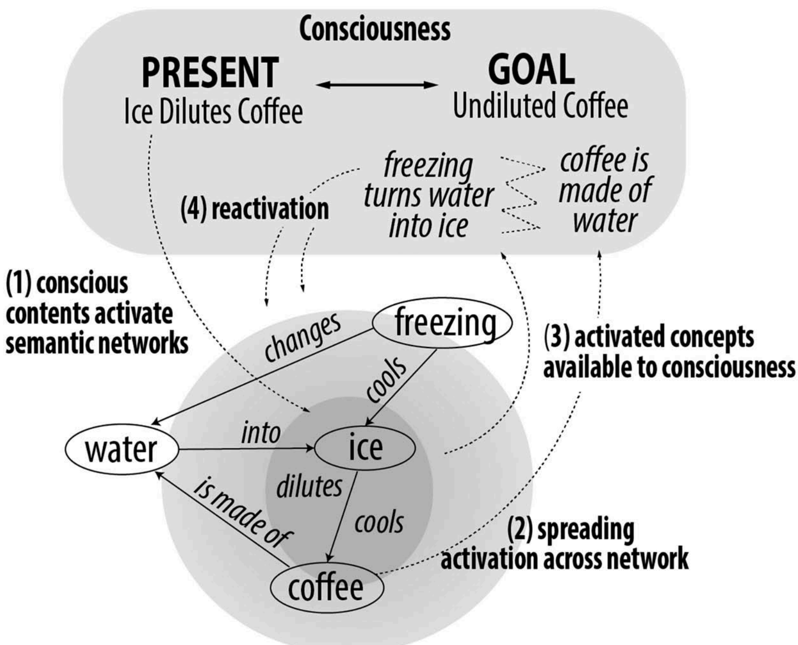

Figure 3. Conscious-unconscious interplay in the construction of novel meaning.

(Bickhard, 2005; Mascolo, Basseches, & El-Hashem, 2014; Zahavi, 2006). Figure 4 extends the embodied model of action to include the social origins of self-awareness. Infants enter the world able to direct their action and attention (Point A) toward both physical and psychological objects (Point B) as well as other people (Point C). As is the case in Figure 3, the arrow at Point A represents the intentionality or directedness of conscious action (*A*) toward its objects (*B*). Although their attention is typically directed outward toward the world, infants also experience a pre-reflective, phenomenal background of bodily and emotional experience (Emde, 1983; Zahavi, 2006). Pre-reflective experience is part of the process of action itself, indicated at Point A. Infants enter the world with a primordial capacity to regulate simple actions (Point 1) and to engage in emotionally regulated *intersubjective* exchanges (Point 2) with others (Reddy, 2015). Intersubjectivity refers to the process by which individual actors share and coordinate experience and meaning between themselves (Mascolo, 2016, 2017; Matusov, 1996).

Although an implicit capacity for self-experience is present in young infants, self-experience changes dramatically over the first two years of life (Rochat, 2013). Around 15–18 months of age, infants gain the capacity to use signs and symbols (i.e., Point 3 in Figure 4) to represent the physical and social world independently of their direct sensorimotor experience. During this time, a *secondary* reflective sense of self emerges as *consciousness takes itself as its own object* (Point 4) (Mead, 1934; Zahavi, 2006). This occurs as consciousness loops back upon itself to form secondary reflexive awareness at Point D. The product of the reflexive looping of consciousness back onto itself is the *secondary sense of self* as an explicit *object of awareness* (Point 4). Over the course of the second year of life, the development of reflexive self-awareness becomes increasingly apparent in the development of self-conscious

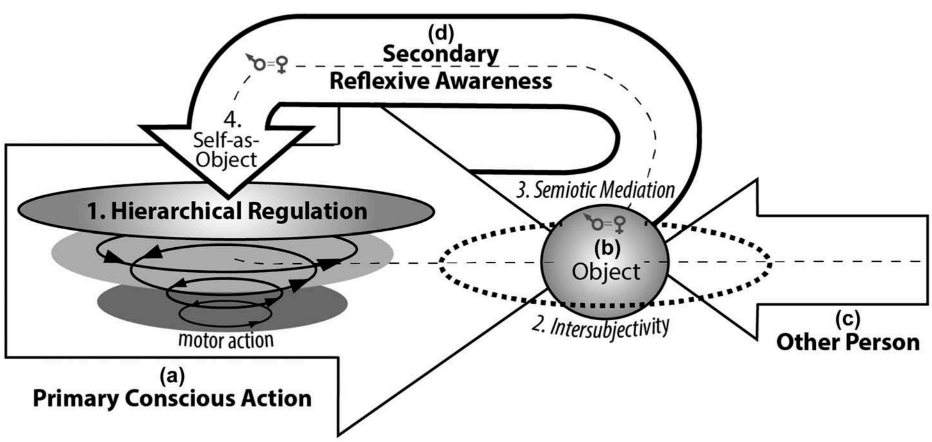

Figure 4. How self-consciousness transforms conscious agency.

emotions like pride and shame (Mascolo & Fischer, 1995), self-recognition in a mirror (Lewis & Carmody, 2008), use of first-person pronouns (Meissner, 2008), and other abilities (Fasig, 2000).

Sign activity (Point 3 in Figure 4) plays a particularly important role in the construction of self. Signs (words) carry the capacity to represent shared social meanings that have sociocultural origins outside of any individual person. It is through the use of signs in discursive interactions with others (Point 2) that children acquire social meanings which are shared within local and extended communities. As children learn to use language, it exerts a *regulatory* (Points 1 and 4) function over thinking, feeling, and motor action (Tartas, Perret-Clermont, & Baucal, 2016). Children construct self-conceptions by appropriating meanings represented in shared and contested conceptions of personhood. When this occurs, *valued images of self come to function as top-level goals and reference standards* (Points 1 and 4) within the hierarchical control systems that drive social action (Carver et al., 2015; Masicampo & Baumeister, 2013). For example, over time, children acquire an understanding of shared and contested representations of gender (Tibbals, 2007). Representations of gender function as top-level goals which drive social action: "I am a boy/girl. This is what it means to be a boy/girl. I will do boy/girl-like things" (Warin, 2000). Thus, higher-order sources of control arise as sign-mediated consciousness gains the capacity to take itself as its own object.

6. What does it mean to be a conscious agent?

The everyday view of conscious agency represents "the person" as a kind of epicenter of action, an underdetermined seat of conscious volition – the originator of actions that are "up to me." These concepts seem central to our sense of conscious agency. Paradoxically, however, perhaps the best way to propose a plausible account of conscious agency is to abandon them. In an embodied approach, it makes sense to say that individuals make choices, exert control, and act on the basis of conscious goals and meanings – but only if we abandon the idea that these capacities arise from a kind of inner or autonomous mental agency. We offer an embodied model as a framework for understanding how persons, as biological organisms, exhibit many, but not all, properties typically attributed to "free will" (Stephan, 2010). The embodied model offers ways of understanding the role of consciousness in orienting the construction of action. It points to ways in which consciousness constrains the operation of neural networks that function to construct *genuinely novel* forms of representation and action. However, it is here that the powers of agency granted by the embodied approach wane. While consciousness contributes to the production of novel action, it cannot do so *autonomously*. While conscious contents activate nonconscious constructive activity, consciousness is itself an emergent product of nonconscious processes. While the endogenous processes that make up an individual person self-organize over time, they do so as parts of larger biological, ecological, social, and cultural systems.

What does it mean, therefore, to speak of the agentive "I"? Shoemaker (1968) suggests that "the word 'I' serves the function of identifying for the audience the subject to which the predicate of the statement must apply" (p. 55). This definition identifies "I" in terms of how it is used within discursive action. However, its psychological meaning hinges upon what it means to speak of a "subject." An embodied view acknowledges at least two uses of "I." The first refers to the capacities of the *person as a whole* to exert control over action – to the person as a unified system whose conscious and unconscious components operate together to produce action. In this sense, "I" refers to "the entirety of this person who appears before you." The second use involves statements that imply a role of *conscious intention* in producing action. If consciousness plays a role in *orienting* and *activating* action, statements invoking the use of "I" to refer to the agentic role of consciousness are meaningful. To say "I will vote for X" identifies the conscious intentions of the speaker as part of the causal structure of action. This is a meaningful statement. It is neither helpful nor necessary, however, to think of agency as something beyond the emergent intention itself – that is, to think of the term "I" as referring to an autonomous self that operates above and beyond the emergent functioning of the organism as a whole.

7. Conclusion

We have sought to offer an embodied and robust conception of conscious agency which avoids the troublesome implications of the notion of *free will*. While most philosophers and psychologists reject mind-body dualism, it is the way we conceptualize the distinction between the mental and the physical that makes the problem of "mental causation" so intractable (Mudrik & Maoz, 2015). The idea of *embodied* emergence can help change how we think about human agency. Instead of seeking to understand how a mental entity with novel causal powers emerges from biophysical processes, it might be better to ask how existing capacities for biological control are transformed by the embodied emergence of meaning and experience (Hateren, 2015). Eschewing free will need not turn us into mere mechanisms or moist robots (Dennett, 2013). We do not need the concept of *free will* to have a robust conception of human agency. As biological systems, humans are not merely moist; they are also conscious. The ability to consciously represent our circumstances transforms biological regulation and is part of what makes us human.

Notes

1. As described by Boogerd et al. (2005), consider the functioning of a cell (*A*) containing an embedded modular system (e.g., an organelle). Such a nested system is composed of the organelle (A_2) and the remainder of the cell in which it is embedded (A_1). Given measures of enzymatic metabolism (e.g., the concentration enzyme metabolites) in both A_1 and A_2, it is possible to measure metabolic output of both A_1 and A_2 as they function in isolation (in test tubes) and as they operate in interaction with each other as part of the larger system itself (*A*). In studies of cell metabolism using *E. coli*, Boogerd et al. (2005) have shown that the metabolic functioning of the cells as a whole (*A*) is often a nonlinear function of the behavior of their components in isolation (A_1, A_2). Boogerd et al. (2005) take the finding that cellular subsystems can exhibit qualitatively different behaviors in isolation than as parts of a larger system as evidence for the *irreducibility* of the systemic properties of the cell as a whole (see Theurer, 2014, for an alternative perspective).
2. To illustrate the concept of hierarchical integration, consider the process by which relational concepts (e.g., cause/effect, reciprocity, temporality, and part/whole) are formed over the course of psychological development (Mascolo & Fischer, 2015; Piaget, 1985; Siegler & Jenkins, 1989). Beginning around two years of age, children gain the capacity able to hold in mind *single concrete thoughts* – symbolic representations that are the equivalent of simple declarative sentences (e.g., "Jack fell down," "He hit me," "Eating candy is fun"). It is not until about three and a half years of age that children gain the capacity to *integrate* two or more such concrete thoughts into a seamless, relational structure – what Fischer (Mascolo & Fischer, 2015) calls "representational mappings." In so doing, children cannot only hold in mind two or more concrete ideas simultaneously; they can also represent the concrete *relationship* between the ideas represented. In so doing, they can represent relationships such as *cause and effect* (e.g., "Jack fell *because* Jill pushed him"), *reciprocity* (e.g., "He hit me *so* I hit him *back*"), *time* (e.g., "We went to the store, and *then* I got candy"), and so forth. Relational concepts thus emerge as higher-order *integrations* of lower-order meanings. The meanings represented in higher-order structures *extend beyond* those contained in their elements, whether those elements are considered alone or in combination. The *reciprocity* communicated in the statement "You hit me, *so* I hit you *back*" goes beyond that communicated by the combination of "You hit me" and "I hit you." Thus, a novel higher-order structure is the emergent *equivalent* of the integration of lower-order elements. There is not (a) the original elements, (b) their integration, *and then also* (c) a novel higher-order structure that exists somehow separate from the lower-order elements. While the novel higher-order structure is *fully dependent* upon the lower-order elements that compose it, it neither exists *separate from* those elements nor is it *reducible to* those elements.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Michael F. Mascolo is Professor of Psychology and Academic Director of the Compass Program at Merrimack College.

Eeva Kallio is Adjunct Professor and Research Associate at University of Jyväskylä and the University of Tempere, Finland.

ORCID

Eeva Kallio http://orcid.org/0000-0003-3950-6554

References

Agafonov, A., & Shabad, S. (2010). Priming effect as a result of the nonconscious activity of consciousness. *Journal of Russian & East European Psychology*, *48*(3), 17–32.

Anderson, P. W. (1972). More is different. *Science*, *177*, 93–396.

Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks*, *20*, 955–961.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer Jr. (Ed.), *The automaticity of everyday life: Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*, 167–203.

Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science*, *41*, 321–333.

Bedau, M. A. (2008). Downward causality and autonomy in weak emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence: Contemporary readings in philosophy and science* (pp. 155–188). Cambridge, MA: MIT Press.

Bich, L., & Moreno, A. (2016). The role of regulation in the origin and synthetic modeling of minimal cognition. *Biosystems,* 12–21.

Bich, L., Mossio, M., Ruiz-Mirazo, K., & Moreno, A. (2016). Biological regulation: Controlling the system from within. *Biology and Philosophy*, *31*(2), 237–265.

Bickhard, M. H. (2005). Consciousness and reflective consciousness. *Philosophical Psychology*, *18*, 205–218.

Bola, M., & Sabel, B. A. (2015). Dynamic reorganization of brain functional networks during cognition. *Neuroimage*, *114*, 398–441.

Bonn, G. B. (2013). Re-conceptualizing free will for the 21st century: Acting independently with a limited role for consciousness. *Frontiers in Psychology*, *4*. Article 920.

Boogerd, F. C., Bruggeman, J., Richardson, R. C., Stephan, A., & Westerhoff, H. V. (2005). Emergence and its place in nature: A case study of biochemical networks. *Synthese*, *145*, 131–164.

Brentano, F. (1874/1995). *Psychology from an empirical standpoint*. New York, NY: Routledge.

Campbell, R. J., & Bickhard, M. H. (2011). Physicalism, emergence and downward causation. *Axiomathes: an International Journal in Ontology and Cognitive Systems*, *21*, 33–56.

Carver, C. S., Johnson, S. L., Joormann, J., & Scheier, M. F. (2015). An evolving view of the structure of self-regulation. In G.H. E. Gendolla, M. Tops, & S. L. Koole (Eds.), *Handbook of biobehavioral approaches to self-regulation* (pp. 9–23). New York, NY: Springer Science + Business Media.

Chalmers, D. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The re-emergence of emergence* (pp. 244–254). Oxford: Oxford University Press.

Charlton, S. G., & Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. Transportation Research Part F. *Traffic Psychology and Behaviour*, *14*(6), 456–471.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. Cambridge, MA: MIT Press.
Dennett, D. C. (2013). *Intuition pumps and other tools for thinking*. New York: Norton.
Di Paolo, E. A., Barandiaran, X. E., Beaton, M., & Buhrmann, T. (2014). Learning to perceive in the sensorimotor approach: Piaget's theory of equilibration interpreted dynamically. *Frontiers in Human Neuroscience, 8*.
Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York, NY: Basic Books.
Ekstrom, L. W. (2000). *Free will: A philosophical study*. Boulder: Westview Press.
Emde, R. N. (1983). The prerepresentational self and its affective core. *The Psychoanalytic Study of the Child*, *38*, 165–192.
Fasig, L. G. (2000). Toddlers' understanding of ownership: Implications for self-concept development. *Social Development*, *9*, 370–382.
Fiacconi, C. M., Dekraker, J., & Köhler, S. (2015). Psychophysiological evidence for the role of emotion in adaptive memory. *Journal of Experimental Psychology: General*, *144*, 925–933.
Fischer, J. M. (1994). *The metaphysics of free will*. Oxford: Blackwell.
Fontaine, J. J. R., & Scherer, K. R. (2013). Emotion is for doing: The action tendency component. In J. J. R. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 170–185). New York, NY: Oxford University Press.
Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, *86*, 829–839.
Freeman, W. J. (2000). Emotion is essential to all intentional behaviors. In M. D. Lewis & I. Granic (Eds.), *Cambridge studies in social and emotional development. Emotion, development, and self-organization: Dynamic systems approaches to emotional development* (pp. 209–235). New York, NY: Cambridge University Press.
Gorniak, S. L., Zatsiorsky, V. M., & Latash, M. L. (2009). Hierarchical control of static prehension: I. Biomechanics. *Experimental Brain Research*, *193*(4), 615–631.
Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*, *37*, 1–47.
Halász, V., & Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sciences*, *2*(2), 130–146.
Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, *8*, 195–207.
Hateren, J. H. (2015). The origin of agency, consciousness, and free will. *Phenomenology and the Cognitive Sciences*, *14*, 979–1000.
Herzog, E. D. (2007). Neurons and networks in daily rhythms. *Nature Reviews Neuroscience*, *8*, 790–802.
Hoel, E. P., Albantakis, L., Marshall, W., & Tononi, G. (2016). Can the macro beat the micro? Integrated information across spatiotemporal scales. *Neuroscience of Consciousness*, *1*. doi:10.1093/nc/niw012
Hommel, B. (2007). Consciousness and control: Not identical twins. *Journal of Consciousness Studies*, *14*, 155–176.
Hume, D. (1748/2008). *An enquiry concerning human understanding*. London, UK: Oxford World Classics.
Jarymowicz, M. (2010). Emotions as evaluative processes: From primary affects to appraisals based on deliberative thinking. In A. Błachnio & A. Przepiórka (Eds.), *Closer to emotions III* (pp. 55–72). Lublin, Poland: Wydawnictwo.

Kallestrup, J. (2006). The causal exclusion argument. *Philosophical Studies: an International Journal for Philosophy in the Analytic Tradition*, *131*, 459–485.

Kallio, E. (2011). Integrative thinking is the key: An evaluation of current research into the development of adult thinking. *Theory & Psychology*, *21*, 785–801.

Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.

Kelso, S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kim, J. (1998). *Mind in a physical world*. Cambridge, MA: MIT Press.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, *95*, 3–36.

Lamme, V. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*, 12–18.

Lamy, D., Leber, A. B., & Egeth, H. E. (2013). Selective attention. In A. F. Healy, R. W. Proctor, I. B. Weiner, A. F. Healy, R. W. Proctor, & I. B. Weiner (Eds.), *Handbook of psychology: Experimental psychology* (Vol. 4, 2nd ed., pp. 267–294). Hoboken, NJ: John Wiley & Sons Inc.

Lewis, M., & Carmody, D. P. (2008). Self-representation and brain development. *Developmental Psychology*, *44*, 1329–1334.

Lewis, M. D. (1996). Self-organising cognitive appraisals. *Cognition and Emotion*, *10*(1), 1–25.

Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, *6*, 47–57.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of freely voluntary act. *Brain*, *106*, 623–642.

Mascolo, M. F., & Fischer, K. W. (1995). Developmental transformations in appraisals for pride, guilt and shame. In K. W. Fischer & J. P. Tangney (Eds.), *Self-conscious emotions: Shame, guilt, embarrassment and pride* (pp. 64–113). New York: Guilford.

Mascolo, M. F., Fischer, K. W., & Li, J. (2003). Dynamic development of component systems of emotions: Pride, shame and guilt in China and the United States. In R. J. Davidson, K. R. Scherer, H. H. Goldsmith, R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 375–408). New York, NY: Oxford University Press.

Mascolo, M. F., Basseches, M., & El-Hashem, A. (2014). Toward an integrated constructionist psychology: The coactive construction of meaning. In J. Raskin, S. Bridges, & J. Kahn (Eds.), *Studies in meaning 5* (pp. 248–300). New York: Pace University Press.

Mascolo, M. F., & Fischer, K. W. (2015). Dynamic development of thinking, feeling and acting. In W. Overton & P. Molennar (Eds.), *Handbook of child psychology and developmental science (vol 1: Theory and method)* (pp. 113–161). New York: John Wiley.

Mascolo, M. F. (2016). Beyond subjectivity and objectivity: The intersubjective foundations of psychological science. *Integrative Psychological & Behavioral Science*, *50*, 185–195.

Mascolo, M. F. (2017). How objectivity undermines the study of personhood: Toward an intersubjective epistemology for psychological science. *New Ideas in Psychology*, *33*, 41–48.

Masicampo, E. J., & Baumeister, R. F. (2011). Finding positive value in human consciousness: Conscious thought serves participation in society and culture. In K. M. Sheldon, T. B. Kashdan, & M. F. Steger (Eds.), Designing positive psychology: Taking stock and moving forward (pp. 175–189). New York, NY: Oxford University Press.

Masicampo, E. J., & Baumeister, R. F. (2013). Conscious thought does not guide moment-to-moment actions—It serves social and cultural functions. *Frontiers in Psychology*, *4*.

Matusov, E. (1996). Intersubjectivity without agreement. *Mind, Culture, and Activity*, *3*(1), 25–45.
Mead, G. H. (1934). *Mind, self and society*. Chicago, IL: Chicago University Press.
Medeiros-Ward, N., Cooper, J. M., & Strayer, D. L. (2014). Hierarchical control and driving. *Journal of Experimental Psychology: General*, *143*(3), 953–958.
Meissner, W. W. (2008). The role of language in the development of the self III: The significance of pronouns. *Psychoanalytic Psychology*, *25*(2), 242–256.
Mele, A. (2009). *Effective intentions: The power of conscious will*. New York, NY: Oxford University Press.
Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about people's folk concept of free will. *Review of Philosophy and Psychology*, *1*, 211–224.
Moors, A., Ellsworth, P., Scherer, K., & Frijda, N. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*, 119–124.
Morsella, E. (2005). The function of phenomenal states: Supramodular Interaction Theory. *Psychological Review*, *112*, 1000–1021.
Morsella, E., Godwin, C. A., Jantz, T. K., Krieger, S. C., & Gazzaley, A. (2015). Homing in on consciousness in the nervous system: An action-based synthesis. *Behavioral and Brain Sciences*.
Morsella, E., Godwin, C. A., Jantz, T. K., Krieger, S. C., & Gazzaley, A. (2017). Homing in on consciousness in the nervous system: An action-based synthesis. *Behavioral and Brain Sciences*, *39*.
Mudrik, L., Deouell, L. Y., & Lamy, D. (2012). Novelty, not integration: Finding the function of conscious awareness. In S. Kreitler, O. Maimon, S. Kreitler, & O. Maimon (Eds.), *Consciousness: Its nature and functions* (pp. 265–276). Hauppauge, NY: Nova Science Publishers.
Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration without awareness: Expanding the limits of unconscious processing. *Psychological Science*, *22*, 764–770.
Mudrik, L., & Maoz, U. (2015). Me & my brain': Exposing neuroscience's closet dualism. *Journal of Cognitive Neuroscience*, *27*, 211–221.
Pacherie, E. (2014). Can conscious agency be saved? *Topoi: An International Review of Philosophy*, *33*(1), 33–45.
Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
Piaget, J. (1985). *The equilibration of cognitive structures*. Chicago, IL: University of Chicago Press.
Pockett, S. (2004). Does consciousness cause behaviour? *Journal of Consciousness Studies*, *11*, 23–40.
Powers, W. T. (2005). *Behavior: The control of perception* (2nd ed.). New Canaan, CT: Benchmark Press.
Prosser, S. (2012). Emergent causation. *Philosophical Studies: an International Journal for Philosophy in the Analytic Tradition*, *159*, 21–39.
Reddy, V. (2015). Joining intentions in infancy. *Journal of Consciousness Studies*, *22*, 24–44.
Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology*, *63*, 101–128.
Rochat, P. (2013). Self-conceptualizing in development. In D. Zelazo & P. D. Zelazo (Eds.), *The Oxford handbook of developmental psychology, vol. 2: Self and other* (pp. 378–397). New York, NY: Oxford University Press.
Roseman, I. J. (1991). Appraisal determinants of discrete emotions. *Cognition and Emotion*, *5*, 161–200.

Roseman, I. J., Wiest, C., & Swartz, T. S. (1994). Phenomenology, behaviors, and goals differentiate discrete emotions. *Journal of Personality and Social Psychology*, *67*, 206–221.
Ruys, K., Stapel, D., & Aarts, H. (2011). From (unconscious) perception to emotion: A global-to-specific unfolding view of emotional responding. In I. Nyklíček, A. Vingerhoets, M. Zeelenberg (Eds.), *Emotion regulation and well-being* (pp. 49–65). New York, NY,US: Springer Science + Business Media.
Sartenaer, O. (2016). Sixteen years later: Making sense of emergence (again). *Journal for General Philosophy of Science*, *47*, 79–103.
Scherer, K. R. (2004). Feelings integrate the central representation of appraisal-driven response organization in emotion. In A. S. R. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions: The Amsterdam symposium* (pp. 136–157). New York, NY: Cambridge University Press.
Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P. U., & Wheatley, T. (2013). Barking up the wrong free: Readiness potentials reflect processes independent of conscious will. *Experimental Brain Research*, *229*, 329–335.
Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy*, *65*, 555–567.
Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Slors, M. (2015). Conscious intending as self-programming. *Philosophical Psychology*, *28*, 94–113.
Smart, J. J. C. (1959). Sensations and brain processes. *The Journal of Philosophy*, *68*, 141–156.
Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543–545.
Sperry, R. W. (1976). Changing concepts of consciousness and free will. *Perspectives in Biology and Medicine*, *20*, 9–19.
Sperry, R. W. (1980). Mind-brain interaction: Mentalism, yes; dualism, no. *Neuroscience*, *5*, 195–206.
Sperry, R. W. (1991). In defense of mentalism and emergent interaction. *Journal of Mind and Behavior*, 12, 221–245.
Stephan, A. (2004). Phenomenal emergence. *Networks*, *3*, 91–102.
Stephan, A. (2010). An emergentist's perspective on the problem of free will. In C. MacDonald & G. MacDonald (Eds.), *Emergence in mind* (pp. 222–239). Oxford: Oxford University Press.
Tartas, V., Perret-Clermont, A., & Baucal, A. (2016). Experimental micro-histories, private speech and a study of children's learning and cognitive development. *Journal for the Study of Education and Development*, *39*, 772–811.
Theurer, K. L. (2014). Complexity-based theories of emergence: Criticisms and constraints. *International Studies in the Philosophy of Science*, *28*, 277–301.
Thompson, E. (2007). *Mind in life*. Cambridge, MA: Harvard University Press.
Tibbals, C. A. (2007). Doing gender as resistance. *Journal of Contemporary Ethnography*, *36*, 731–751.
Tomkins, S. S. (1984). Affect theory. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 163–195). Hillsdale, NJ: Erlbaum.
Tononi, G. (2007). The information integration theory of consciousness. In M. Velmans, S. Schneider, M. Velmans, & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 287–299). Malden: Blackwell Publishing.

Tononi, G. (2013). On the irreducibility of consciousness and its relevance to free will. In A. Suarez, P. Adams, A. Suarez, & P. Adams (Eds.), *Is science compatible with free will? Exploring free will and consciousness in the light of quantum physics and neuroscience* (pp. 147–176). New York, NY: Springer Science + Business Media.

Warin, J. (2000). The attainment of self-consistency through gender in young children. *Sex Roles*, *42*, 209–231.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Werner, H., & Kaplan, B. (1963/1984). *Symbol formation.* Hillsdale, NJ: Lawrence Erlbaum Publishers.

Wilson, J. (2015). Metaphysical emergence: Weak and strong. In B. Tomasz & C. Wuthrich (Eds.), *Metaphysics in contemporary physics* (pp. 345–402). Amsterdam: Rodopi.

Yamaguchi, M., Crump, M.J. C., & Logan, G. D. (2013). Speed-accuracy trade-off in skilled typewriting: Decomposing the contributions of hierarchical control loops. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 678–699.

Zahavi, D. (2006). *Subjectivity and selfhood*. Cambridge, MA: MIT Press.