Graziano MSA, Morsella E (2020) A new motor approach to consciousness: Implications for the simulation of future behavior. Journal of Consciousness Studies, 27: 88-103.

Michael S.A. Graziano
and Ezequiel Morsella

A New Motor Approach to Consciousness

Implications for the Simulation of Future Behaviour

Abstract: *In the last ten years, several new scientific approaches to consciousness have emerged. Two in particular have a close relationship to each other: passive frame theory and the attention schema theory. Both take movement control as a starting point, a relatively new perspective on the topic, and both emphasize the importance of consciousness for future, anticipated actions. Passive frame theory proposes that consciousness serves as a field of information that enables flexible, context-sensitive action selection for to-be-produced actions. The attention schema theory suggests that consciousness is related to the control mechanisms for attention, which include models of the world and also of the actor's own mental processes. This paper (a) discusses the relationship between the two theories and asks whether they can be considered as facets of the same underlying mechanism, and (b) attempts to illuminate how such processes associated with consciousness are essential for the simulation of future, anticipated actions.*

Correspondence:
Email: graziano@princeton.edu

1. Introduction

Some of the processes in the brain have consciousness[1] attached to them, whereas others do not (Goodale and Milner, 2004; Gray, 2004; Koch *et al.*, 2016). Exactly what makes that difference remains debated (Kinsbourne, 1996). Many cognitive and neuroscientific approaches to consciousness have been proposed, such as the global workspace theory (Baars, 1988; Dehaene, 2014; Newman and Baars, 1993), integrated information theory (Oizumi, Albantakis and Tononi, 2014; Tononi, 2008), and higher-order thought theory (Lau and Rosenthal, 2011; Rosenthal, 2005). In these and other approaches, investigators often focus on perceptual processing with an emphasis on vision. The implication is that if we can understand how one becomes conscious of a visual stimulus, then the solution can be generalized to any internal process, and the essentials of consciousness will be understood.

Here we discuss a different possible approach, focusing on action rather than on perception. We suggest that the brain's control of complex, adaptive action comes with a set of computational requirements that may provide explanations missing from the sensory-only approach. This paper focuses on two recently proposed action-based theories,[2] passive frame theory (Morsella *et al.*, 2016; see Dou, Walker and Morsella, this issue) and the attention schema theory (Graziano, 2013). The purpose of the paper is to explore how two seemingly different theories may complement each other in constructive ways. Our hope is to provide an example of inter-theory diplomacy — how theory building can proceed by cross-fertilization between developed frameworks. We first summarize each theory and then discuss the possible relationship between them. Second, we

1 Nagel (1974) claimed that an organism possesses consciousness if there is *something it is like* to be that organism — something it is like, for example, to be human and experience pain, breathlessness, or yellow after-images. Block (1995) similarly claimed, 'the phenomenally conscious aspect of a state is what it is like to be in that state' (p. 227).

2 The two theories presented in this paper are two of a handful of action-based theories of consciousness, including the sensorimotor account by O'Regan and Noë (2001) and theories in which the nature of percepts is based primarily on motor processing, as in 'peripheralist', 'motor', 'efferent', and 'reafferent' theories of thought (e.g. Festinger *et al.*, 1967; Hebb, 1968; Held and Rekosh, 1963; McGuigan, 1966; Münsterberg, 1891; Washburn, 1928; Watson, 1924). See Scheerer (1984) for a review of the shortcomings of such approaches.

attempt to illuminate how processes associated with consciousness are essential for the simulation of future, anticipated actions.

2. Passive Frame Theory (PFT)

The main message of PFT is that consciousness contributes to complex, adaptive action. When people select an action, that selection can be influenced by many different sources of information. According to PFT, integrating those sources of information for the purposes of controlling behaviour at each moment in time requires consciousness.

Imagine you are underwater and have the urge to inhale but, because of the specific circumstances, you suppress the urge. Or imagine you are carrying a hot dish of food and have an urge to drop the dish, yet refrain from doing so. In these examples, the action that is selected depends on context. We call this 'integrated' behaviour because it is selected on the basis of integrating contextual information. In PFT, without consciousness, sophisticated behaviour can still arise but will not be integrated. For example, in neurological conditions in which actions are decoupled from consciousness and arise involuntarily, complex actions such as manipulating tools or removing clothing can arise, but the actions are not influenced by the kinds of information that should normally influence them. In that sense, these actions are 'unintegrated', appearing insensitive to context. Thus, motor control and sophisticated perceptual processing can occur unconsciously, but flexibly using perceptual representations to shape action selection cannot occur without consciousness.

In this approach, consciousness contains a field of information content. The behavioural response to any one item of content in that field can be 'framed' by the other contents composing the field at that time. This integrative process is called 'collective influence' (Morsella, 2005; Morsella *et al.*, 2016) and is essential for adaptive, to-be-produced behaviours. The conscious field thus permits the collective influence of all *conscious contents*[3] activated at a given time. When consciousness fails, overt action is not 'integrated' in this way. In anarchic hand syndrome, for example, the hand might grab an object that belongs to someone else or might, out of the blue, unbutton a

[3] A 'conscious content' is any mental representation of which one is aware (Merker, 2007). For example, it might be a colour, an urge, or a spontaneous memory. The 'conscious field' is all that one is aware of at one moment in time, which is the combination of all activated conscious contents (Freeman, 2004; Köhler, 1947; Searle, 2000).

button on the sleeve. These actions are not unsophisticated behaviours (consider that neither a robot nor a 3-year-old can, for instance, unbutton clothing). Rather, these actions are 'unintegrated' actions.

According to PFT, the conscious field includes potential 'action options' (see Dou, Walker and Morsella, this issue). These options include percepts, urges, memories, and mental imagery. For example, one may act out or suppress a consciously experienced action-related urge, or one may act toward a stimulus that is nearby on the left versus an object that is further away on the right. These spatial aspects of action selection require the creation of a first-person perspective (Merker, 2013), which is a basic part of conscious experience, both while awake and even in the dream world. Thus, from the requirements of action selection, a primitive form of self must emerge. It is important to clarify that this form of self is not that which the sense of self is commonly understood to be. Instead, it is the first-person perspective that is immanent in perceptual experience.

If PFT is correct, then the role of consciousness is more circumscribed than previously thought. Consciousness serves the somatic nervous system, which controls skeletal muscle output. Even within the domain of skeletal movements, consciousness is limited: the mechanisms that generate the contents of consciousness, and the mechanisms that 'sample' the conscious field and generate actions, are themselves unconscious. One conscious content does not, in a sense, 'know' of the nature of other conscious contents nor of the nature of ongoing behaviour. This peculiar 'insulated' property of conscious contents is proposed to be evolutionarily adaptive (see Morsella *et al.*, 2016, for further discussion).

In PFT, consciousness is much more passive and less purposeful than in other accounts. Consciousness has no reasoning, no memory, or symbol manipulation of its own. Consciousness does the same thing (that is, to permit collective influence) over and over, for various kinds of processes, making it seem that it is more flexible than it is and that it can do more things than it does. Although consciousness is passive, what it provides to action systems is essential for adaptive behaviour, just as a window is passive but essential for the driver of a car.

Consistent with 'global workspace' approaches (Baars, 1988; 2002; Dehaene, 2014), PFT proposes that these states integrate nervous processes that are otherwise independent. However, unlike the workspace models (e.g. Baars, 1988; Dehaene, 2014), which propose that conscious representations are broadcast to modules engaged in both

stimulus interpretation and content generation, in PFT (as in Merker, 2007) the contents of the conscious field are directed only at the unconscious processes of the skeletomotor output system. In addition, in workspace approaches, consciousness serves more than a handful of functions (e.g. Baars, 1988; Dehaene, 2014), including adaptation and learning, decision making, analogy forming, editing and debugging, metacognitive self-monitoring, and autoprogramming (Baars, 1988). In PFT, the conscious field serves only one basic, passive role.

3. The Attention Schema Theory (AST)

The AST grew out of a consideration of movement control and the role of the body schema, the brain's internal model of the physical body. A machine computes a model of a part of itself in order to better control that part of itself (Camacho and Bordons Alba, 2004). That reliance on an internal model is a basic principle of dynamical systems control. The models do not need to be exact or detailed. For example, the body schema is not a microscopically exact description of the body. It does not model bone structure or muscle fibre composition. The body schema is more like a cartoon sketch, an approximation. And yet it is necessary for good control of the body. It allows the control system to monitor the basic structure of the body, how it is moving each moment, and how it is likely to move in the next few moments (Graziano and Botvinick, 2002; Scheidt *et al.*, 2005; Wolpert, Ghahramani and Jordan, 1995). When the body schema fails, movement control is possible but is much less effective and suffers inaccuracies and instabilities.

In the AST, the brain controls more than just the movement of the body. It also controls a different kind of movement — the virtual movement of attention (Graziano, 2013; Webb and Graziano, 2015; Webb, Kean and Graziano, 2016). By attention, we refer to the selective enhancement of some signals over others and the deep processing of those enhanced signals (Beck and Kastner, 2009; Desimone and Duncan, 1995). Selective attention is constantly moving from one set of items to another, for example, from the words at your central gaze, to a sound in your back yard, to a memory you have just recalled, to an emotion or a thought. That movement of attention is not random — it is strategic and controlled. According to the principles of control theory, in order to control attention efficiently and accurately, the brain must construct an internal model of it — an attention schema — a constantly updated, approximate representation of what attention is,

what it does, what its consequences are, and what state attention is in at each moment.

Imagine a person is looking at an apple. In this theory, the brain constructs three key representations. First, it constructs a representation of the apple, perhaps mainly in the visual cortex. The representation of the apple allows the brain to arrive at the conclusion: 'There is an apple. It's red, round, and located over there.' Second, the brain constructs a complex, rich representation of the self, information about one's own body, location, and personal history. The self model allows the brain to conclude: 'There is a me. I am a person. I'm standing right here. I walked here.' And yet, with only these two representations, the brain still lacks sufficient information to conclude or report anything about consciousness of the apple.

Third, crucially, the brain constructs a representation of the rich, complex process of attention. Accessing the information in that model of attention, the brain can conclude: 'I am conscious of the apple. By that I mean not only is there an apple, and not only is there a me, but I have a mental possession of the apple. It takes up some of my mental resources. I am able to respond to the apple in several ways and can even remember it so that I can choose to respond later.' According to the AST, people assert that they have consciousness because the brain has access to the information in an attention schema. Without an attention schema, the brain would have no concept of consciousness. It would have no basis on which to answer questions about it. The very construct of consciousness — of a mind that has a personal grasp of something and can thereby choose to react to it — is a simplified model of attention. Unlike higher-order thought theory, the AST focuses on how consciousness can emerge, in the context of perception-and-action, from the internal modelling of, specifically, attention.

According to the AST, people think of consciousness as a non-physical or metaphysical mystery because, just as the body schema is only a caricature of what it represents, the attention schema, too, provides only a caricature of what it represents. It describes attention in a blurred manner, as a metaphysical force without the mechanistic details of how attention is implemented at the level of neurons and synapses.

4. PFT and the AST, Side by Side

Figure 1 illustrates the possible relationship between PFT and the AST. The AST deals in how items are selected to enter a field of attended items and why we attach the property of consciousness to that field. PFT deals in how actions are selected, based on the influence of items in that conscious field.

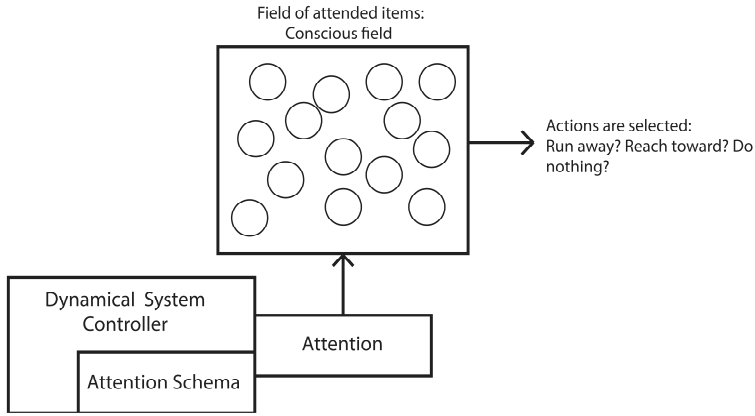

Figure 1. Passive frame theory and the attention schema theory working together. In the attention schema theory, attention is controlled by a dynamical system controller. One part of that controller is an attention schema. The information contained in the attention schema forms the basis from which we claim to have a conscious experience of the attended items. Items selected by attention are shown in the figure as part of a 'conscious field'. In PFT, the purpose of that conscious field is to flexibly influence action selection.

It has long been recognized that attention has a close relationship to movement control. After all, the more attention a stimulus receives, the more likely it is to influence action. Suppose an apple sits in front of you. The more attentional enhancement your brain applies to that visual signal, the more likely you are to select an action toward it. Conversely, you are very unlikely to react to the apple if you are directing no attention to it. Thus, although attention has classically been considered a sensory phenomenon and has been studied mainly in the visual domain, it is inextricably linked to action control (Allport, 1989; Moore, Armstrong and Fallah, 2003; Neumann, 1987; Noudoost *et al.*, 2010; Rizzolatti, 1983; Rizzolatti *et al.*, 1987). Not only does attention to an item make you more likely to look at it and reach toward it, but looking at it and reaching toward it enhances your

attention to it: one of the primary top-down signals that tips the perceptual competition toward one or another stimulus is the feedback signal coming from movement planning regions of the brain, especially those for the planning of eye movements (Moore, Armstrong and Fallah, 2003; Noudoost *et al*., 2010).

The close link between attention and action selection should not be surprising. An animal can choose only one direction to walk at a time, chew only one mouthful at a time, direct the eyes and head toward one object at a time, and grasp and manipulate only a few items at a time (Neumann, 1987). The system therefore evolved to focus its sensory resources on one part of the world at a time. This close working relationship between attentional selection and action selection suggests that there may be an equally close relationship between PFT and the AST.

In PFT, the conscious field is composed of a medley of action options and items that can influence action selection. Moreover, the content in this conscious field is represented from a first-person perspective, since that is the format required for action control. But why is subjectivity a part of this multi-determined action-selection process? Why can't a machine select actions without also claiming to possess consciousness?

At each moment, a person is confronted by many different aspects of the world and many different movement options. A mechanism selects only a subset of that content to enter the field illustrated in Figure 1. According to PFT, only those selected, enhanced items can collectively influence action selection, for adaptive, future behaviour. In the AST, the attentional mechanism selecting that subset employs standard principles of control, including an attention schema. It is the information in the attention schema that forms the basis on which we claim to have consciousness of the attended items. Thus, the two theories tackle the problem of consciousness from complementary vantage points, with PFT focusing on the importance of consciousness for adaptive action selection and the AST emphasizing how the mechanisms underlying attention cause us to claim to have a conscious experience.

4.1. Similarities between the two approaches

1. In addition to focusing on action instead of on perception, both approaches are 'functionalist' (in the tradition of William James): consciousness is for adaptive, to-be-produced behaviour. It is

concerned with immediate, to-be-produced actions. This view is in contrast to other accounts (e.g. Hommel, 2013; Masicampo and Baumeister, 2013; Prinz, 2012).

2. In both approaches, the conscious field represents a simplified and inaccurate model of the world, one in which, in line with functionalist approaches, adaptive action selection trumps representational accuracy. Hence, H_2O_2 (hydrogen peroxide) does not, to one, taste like water (H_2O) with just 'a little extra oxygen' but rather, because of its poisonous properties, it tastes like something that should be expelled from the mouth immediately. Here, the properties of the percept reflect not the way things are, but rather how one should respond to them. Moreover, the contents composing the attended field (e.g. the colour white) do not exist in the physical world; they are as contrived as the icons that, on a radar screen, represent a plane versus a cloud. As Lashley (1923) notes:

> Quality is something unique, indescribable, except in terms of itself. Red is red, green is green. Neither is, by any stretch of the imagination, a form of ether vibration or chemical change in the brain… when by analysis the simplest qualities are reached, nothing more can be said of them save that they are in different, undefinable degrees diverse. They have no describable characters inherent in themselves; they are not analyzable into anything else. They exist by virtue of their indescribable differences and by virtue of nothing else discoverable by introspection. (pp. 252–3)

3. In both approaches, though consciousness is for adaptive behavioural control, the format of conscious representations is 'perceptual-like'. During the control of behaviour, for example, one is unconscious of the efference to the muscles that dictates which fibres should be activated at which time (Rosenbaum, 2002). Although one is unconscious of these complex programs, one is often aware of their proprioceptive and perceptual consequences, such as seeing the hand grasp a cup or hearing oneself utter 'hello'.

4. Both approaches emphasize that conscious processing is intimately linked to the physical limitations of action production. Because of the mechanical limitations of the skeletal muscle effector system, one can perform only one or a few actions at a time (Lashley, 1951; McFarland and Sibly, 1975; Sherrington, 1906). For example, one can utter only one word at a time (Wundt, 1900). Hence, there is the need for a mechanism that somehow prevents responding to all things (e.g. stimuli, urges, and memory-based information) and prioritizes responding to some things over other things (see Tsotsos, 2011).

According to Graziano (2013) and other theorists (Allport, 1989; Neumann, 1987), attention provides this essential function (Krauzlis *et al.*, 2014). Graziano (2013) goes on to reason that if attention itself needs to be controlled, then it is beneficial for there to be an internal model of it, which is the attention schema mentioned above. Second, Graziano reasons that if there is an attention schema, then one would claim to have awareness based on the information in that schema.

4.2. Consciousness and the mental simulation of future actions

The foregoing has implications for our understanding of the role of consciousness in the mental simulation of to-be-produced, anticipated action. For example, according to PFT, when one feels an urge to produce a certain action, this urge cannot be modulated or turned off at will. Moreover, PFT argues that such modulations would not be adaptive in the long term. Hence, in the model, inclinations regarding future behaviour, which can be construed as 'action options', can be behaviourally suppressed, but often they cannot be mentally suppressed (Bargh and Morsella, 2008). The unconscious systems triggering urges cannot (and should not) influence behaviour directly, but they influence the nature of consciousness. This accords well with the view that, for adaptive future action, it is best for urges to function, not as 'programs' that inflexibly and directly trigger action, but rather only as 'advice' for possible actions (Agre and Chapman, 1990; Morsella *et al.*, 2016).

From this standpoint, stimulus-elicited urges function much like the 'internalized reflexes' mentioned by Vygotsky (1962). These internalized reflexes can be co-opted to play an essential role in the mental simulation of future actions (Bargh and Morsella, 2008). Short of performing an action, one good way of learning the potential consequences of a course of action is to simulate the action. This has been long known to engineers. Today, there are robotics that, before responding to a stimulus, run simulations of the possible consequences of responding to the stimulus in one way versus another. One obvious value of simulation is that knowledge of the consequences of a given plan is learned without the risks of performing the actions. Similarly, according to the AST, the brain constructs an internal model, or representation, not only of behaviour, but also of a mental process — attention. This 'attention schema' is used to help control attention, much like the 'body schema', the brain's internal simulation of the body, is used to help control the body and to guide future actions.

From this standpoint, consciousness is the caricature of attention depicted by that internal model.

For such simulations about anticipated behaviours to be useful, they must be based to some extent on solid knowledge regarding both the consequences of an action and the actions by others (anticipated actions) that might follow one's action. In addition, such mental simulations require a separate process in which the *simulacra* (i.e. the products of simulation) are somehow evaluated. To take an example from Bargh and Morsella (2008), if an army general had no clue regarding what constitutes a favourable battle outcome, then there would be no use in simulating battle formations. The point is that simulation can construct simulacra, and these simulacra can be of high or low quality (depending in part on the knowledge on which they are built); however, simulation by itself cannot evaluate the simulacra. Such evaluation is challenging because it depends on taking into account a wide variety of considerations (e.g. physical or social consequences). Conveniently, most knowledge regarding what constitutes favourable or unfavourable outcomes is already available in the mental apparatus: it resides in the unconscious systems that, before the advent of suppression, controlled behaviour directly. These now 'smothered' agents, which produce urges and other behavioural inclinations, respond to simulacra as if they were responding to an actual external stimulus. The internalized reflexes provide the evaluative judgment or 'gut feelings' that mental simulations require. Hence, unconscious processes not only adapt us to the present situation but also influence the tracks we lay to guide our future behaviour. Importantly, the simulation concerns, not the simulation of future mental processes, but of future action.

4.3. The principle of atemporality

According to PFT, a given combination of behavioural inclinations in the conscious field will always, in principle, yield the same outcome regarding action selection (Morsella *et al.*, 2016). This is the case regardless of whether the inclinations are triggered by external stimuli or through mental simulation (Bhangal *et al.*, 2016). From this standpoint, with all things being equal, the combination of inclinations *X*, *Y*, and *Z* will always yield the selection of action *Z* (or, more precisely, of 'operant' *Z*). Thus, one would *always* have a negative gut feeling about saying a funny statement at a serious event or about carrying a heavy vase on a freshly mopped floor. One would have these negative

feelings regardless of whether the acts were to be done in the next five minutes or long into the future. Hence, the same collection of inclinations will always yield the same judgment regarding action selection. This has been referred to as the *principle of atemporality* (Bhangal *et al*., 2016). According to PFT, such an architecture arises because inclinations do not, in a sense, 'know' whether they are relevant to ongoing action or to the other contents composing the conscious field (see Dou, Walker and Morsella, this issue). It is important to note that, according to this principle, the same medley of conscious inclinations will always yield the same intrapsychic conflicts. These conflicts will arise regardless of whether these inclinations are activated by external stimuli, information in working memory (Hubbard *et al.*, 2013), or mental simulations of future, anticipated actions (Bhangal *et al.*, 2016).

Thorndike (1905) concludes, 'The function of thoughts and feelings is to influence actions… Thought aims at knowledge, but with the final aim of using the knowledge to guide action' (p. 111). In accord with Thorndike's conclusion, both the AST and PFT take movement control as a starting point, a relatively new perspective on the topic of consciousness. The two theories fit together as neatly as two lego blocks. At each moment, a person is confronted by many different aspects of the world. An attentional mechanism selects a subset of that content. Those selected, enhanced items can collectively influence action selection. That is the PFT part of the process. The attentional mechanism uses standard principles of control, and therefore it includes an attention schema. The information in that attention schema forms the basis on which we claim to have consciousness of the attended items. That is the AST part of the process. The two theories complement each other: PFT emphasizing the importance of action selection for consciousness and the AST emphasizing how the mechanisms underlying attention cause us to claim to have a conscious experience.

Whether either theory is correct, of course, remains an empirical question. The lesson here is that the theories, as contrasting as they may initially seem, should not be viewed as rivals. They can potentially fit together to offer a deeper understanding.

Many alternative theories of consciousness have been proposed. For example, other theoretical approaches have posited that the integration associated with consciousness is for high-level semantic processes (e.g. Thagard and Stewart, 2014). In other accounts, consciousness is not for intra-organismic processes, but for high-level, sociocultural

interactions (Banks, 1995; Carlson, 1994; Frith, 2010; Macphail, 1998; Prinz, 2012). Some have hypothesized that consciousness serves no function in action control (Hommel, 2013; Koch, 2014; Masicampo and Baumeister, 2013; see also Jackson, 1986; Kinsbourne, 1996; 2000; Pinker, 1997).

Typically proponents of each theory erect defences for the home theory and counter-arguments for the competing theories. Some of these competing theories, however, might have commonalities. When explored in sufficient depth, they may turn out to connect at a deeper level. We offer this review as an example of how disparate theories of consciousness, seemingly in competition with each other, can turn out to complement each other.

List of Recommended Readings

Goodale, M. & Milner, D. (2004) *Sight Unseen: An Exploration of Conscious and Unconscious Vision*, New York: Oxford University Press.

Gray, J. (2004) *Consciousness: Creeping Up on the Hard Problem*, New York: Oxford University Press.

Graziano, M.S.A. (2013) *Consciousness and the Social Brain*, New York: Oxford University Press.

Hommel, B. (2013) Dancing in the dark: No role for consciousness in action control, *Frontiers in Psychology*, **4**, art. 380.

Koch, C. (2004) *The Quest for Consciousness: A Neurobiological Approach*, Englewood, CO: Roberts & Company.

References

Agre, P.E. & Chapman, D. (1990) What are plans for?, *Robotics and Autonomous Systems*, **6**, pp. 17–24.

Allport, D.A. (1989) Visual attention, in Posner, M.I. (ed.) *Foundations of Cognitive Science*, vol. 2, pp. 631–682, Cambridge, MA: MIT Press.

Baars, B.J. (1988) *A Cognitive Theory of Consciousness*, New York: Cambridge University Press.

Baars, B.J. (2002) The conscious access hypothesis: Origins and recent evidence, *Trends in Cognitive Sciences*, **6**, pp. 47–52.

Banks, W.P. (1995) Evidence for consciousness, *Consciousness and Cognition*, **4**, pp. 270–272.

Bargh, J.A. & Morsella, E. (2008) The unconscious mind, *Perspectives on Psychological Science*, **3**, pp. 73–79.

Beck, D.M. & Kastner, S. (2009) Top-down and bottom-up mechanisms in biasing competition in the human brain, *Vision Research*, **49**, pp. 1154–1165.

Bhangal, S., Cho, H., Geisler, M.W. & Morsella, E. (2016) The prospective nature of voluntary action: Insights from the reflexive imagery task, *Review of General Psychology*, **20**, pp. 101–117.

Block, N. (1995) On a confusion about a function of consciousness, *Behavioral and Brain Sciences*, **18**, pp. 227–287.

Camacho, E.F. & Bordons Alba, C. (2004) *Model Predictive Control*, New York: Springer.
Carlson, N.R. (1994) *Physiology of Behavior*, Needham Heights, MA: Allyn and Bacon.
Dehaene, S. (2014) *Consciousness and the Brain*, New York: Viking Press.
Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention, *Annual Review of Neurosciences*, **18,** pp. 193–222.
Dou, W., Walker, E.B. & Morsella, E. (this issue) The prospective nature of involuntary entry into consciousness, *Journal of Consciousness Studies*, **27** (3–4).
Festinger, L., Ono, H., Burnham, C.A. & Bamber, D. (1967) Efference and the conscious experience of perception, *Journal of Experimental Psychology Monograph*, **74**, pp. 1–36.
Freeman, W.J. (2004) William James on consciousness, revisited, *Chaos and Complexity Letters*, **1**, pp. 17–42.
Frith, C.D. (2010) What is consciousness for?, *Pragmatics & Cognition*, **18** (3), pp. 497–551.
Goodale, M. & Milner, D. (2004) *Sight Unseen: An Exploration of Conscious and Unconscious Vision*, New York: Oxford University Press.
Gray, J.A. (2004) *Consciousness: Creeping Up on the Hard Problem*, New York: Oxford University Press.
Graziano, M.S.A. (2013) *Consciousness and the Social Brain*, New York: Oxford University Press.
Graziano, M.S.A. & Botvinick, M.M. (2002) How the brain represents the body: Insights from neurophysiology and psychology, in Prinz, W. & Hommel, B. (eds.) *Common Mechanisms in Perception and Action: Attention and Performance XIX*, pp. 136–157, New York: Oxford University Press.
Hebb, D.O. (1968) Concerning imagery, *Psychological Review*, **75**, pp. 466–477.
Held, R. & Rekosh, J. (1963) Motor-sensory feedback and the geometry of visual space, *Science*, **141**, pp. 722–723.
Hommel, B. (2013) Dancing in the dark: No role for consciousness in action control, *Frontiers in Psychology*, **4**, art. 380.
Hubbard, J., Rigby, T., Godwin, C.A., Gazzaley, A. & Morsella, E. (2013) Representations in working memory yield interference effects found with externally-triggered representations, *Acta Psychologica*, **142**, pp. 127–135.
Jackson, F. (1986) What Mary didn't know, *The Journal of Philosophy*, **83**, pp. 291–295.
Kinsbourne, M. (1996) What qualifies a representation for a role in consciousness?, in Cohen, J.D. & Schooler, J.W. (eds.) *Scientific Approaches to Consciousness*, pp. 335–355, Hillsdale, NJ: Erlbaum.
Kinsbourne, M. (2000) How is consciousness expressed in the cerebral activation manifold?, *Brain and Mind*, **2**, pp. 265–274.
Koch, C. (2014) *Consciousness: Colloquium Delivered at the Redwood Center for Theoretical Neuroscience*, 2 December 2014, University of California, USA.
Koch, C., Massimini, M., Boly, M. & Tononi, G. (2016) Neural correlates of consciousness: Progress and problems, *Nature Reviews Neuroscience*, **17**, pp. 307–321.
Köhler, W. (1947) *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*, New York: Liveright.

Krauzlis, R.J., Bollimunta, A., Arcizet, F. & Wang, L. (2014). Attention as an effect not a cause, *Trends in Cognitive Sciences*, **18**, pp. 457–464.

Lashley, K.S. (1923) The behavioristic interpretation of consciousness, *Psychological Bulletin*, **30**, pp. 237–272 (Part 1), 329–353 (Part 2).

Lashley, K.S. (1951) The problem of serial order in behavior, in Jeffress, L.A. (ed.) *Cerebral Mechanisms in Behavior: The Hixon Symposium*, pp. 112–146, New York: Wiley.

Lau, H. & Rosenthal, D. (2011) Empirical support for higher-order theories of consciousness, *Trends in Cognitive Sciences*, **15**, pp. 365–373.

Macphail, E.M. (1998) *The Evolution of Consciousness*, New York: 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 Psycholology*, **4**, art. 478.

McFarland, D.J. & Sibly, R.M. (1975) The behavioural final common path, *Philosophical Transactions of the Royal Society (London)*, **270**, pp. 265–293.

McGuigan, F.J. (1966) *Thinking: Studies of Covert Language Processes*, New York: Appleton-Century-Crofts.

Merker, B. (2007) Consciousness without a cerebral cortex: A challenge for neuroscience and medicine, *Behavioral and Brain Sciences*, **30**, pp. 63–134.

Merker, B. (2013) The efference cascade, consciousness, and its self: Naturalizing the first person pivot of action control, *Frontiers in Psycholology*, **4**, pp. 1–20.

Moore, T., Armstrong, K.M. & Fallah, M. (2003) Visuomotor origins of covert spatial attention, *Neuron*, **40**, pp. 671–683.

Morsella, E. (2005) The function of phenomenal states: Supramodular interaction theory, *Psychological Review*, **112**, pp. 1000–1021.

Morsella, E., Godwin, C.A., Jantz, T.J., Krieger, S.C. & Gazzaley, A. (2016) Homing in on consciousness in the nervous system: An action-based synthesis, *Behavioral and Brain Sciences*, **39**, pp. 1–17.

Münsterberg, H. (1891) Über Aufgaben und Methoden der Psychologie, *Schriften der Gesellschaft für psychologische Forsschung*, **1**, pp. 93–272.

Nagel, T. (1974) What is it like to be a bat?, *Philosophical Review*, **83**, pp. 435–450.

Neumann, O. (1987) Beyond capacity: A functional view of attention, in Heuer, H. & Sanders, A.F. (eds.) *Perspectives on Perception and Action*, pp. 361–394, Hillsdale, NJ: Lawrence Erlbaum.

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

Noudoost, B., Chang, M.H., Steinmetz, N.A. & Moore, T. (2010) Top-down control of visual attention, *Current Opinions in Neurobiology*, **20**, pp. 183–190.

Oizumi, M., Albantakis, L. & Tononi, G. (2014) From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0, *PLoS Computational Biology*, **10**, e1003588.

O'Regan, J.K. & Noë, A. (2001) A sensorimotor account of vision and visual consciousness, *Behavioral and Brain Sciences*, **24**, pp. 883–917.

Pinker, S. (1997) *How the Mind Works*, New York: Norton.

Prinz, W. (2012) *Open Minds: The Social Making of Agency and Intentionality*, Cambridge, MA: MIT Press.

Rizzolatti, G. (1983) Mechanisms of selective attention in mammals, in Ewert, J.P. (ed.) *Advances in Vertebrate Neuroethology*, pp. 261–297, London: Plenum.
Rizzolatti, G., Riggio, L., Dascola, I. & Umiltá, C. (1987) Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention, *Neuropsychologia*, **25**, pp. 31–40.
Rosenbaum, D.A. (2002) Motor control, in Pashler, H. (series ed.) & Yantis, S. (vol. ed.) *Stevens' Handbook of Experimental Psychology: Vol. 1. Sensation and Perception*, 3rd ed., pp. 315–339, New York: Wiley.
Rosenthal, D.M. (2005) *Consciousness and Mind*, Oxford: Clarendon Press.
Scheerer, E. (1984) Motor theories of cognitive structure: A historical review, in Prinz, W. & Sanders, A.F. (eds.) *Cognition and Motor Processes*, pp. 77–98, Berlin: Springer-Verlag.
Scheidt, R.A., Conditt, M.A., Secco, E.L. & Mussa-Ivaldi, F.A. (2005) Interaction of visual and proprioceptive feedback during adaptation of human reaching movements, *Journal of Neurophysiology*, **93**, pp. 3200–3213.
Searle, J.R. (2000) Consciousness, *Annual Review of Neurosciences*, **23**, pp. 557–578.
Sherrington, C.S. (1906) *The Integrative Action of the Nervous System*, New Haven, CT: Yale University Press.
Thagard, P. & Stewart, T.C. (2014) Two theories of consciousness: Semantic pointer competition vs. information integration, *Consciousness and Cognition*, **30**, pp. 73–90.
Thorndike, E.L. (1905) The functions of mental states, in Thorndike, E.L. (ed.) *The Elements of Psychology*, pp. 111–119, New York: A.G. Seiler.
Tononi, G. (2008) Consciousness as integrated information: A provisional manifesto, *Biological Bulletin*, **215**, pp. 216–242.
Tsotsos, J.K. (2011) *A Computational Perspective on Visual Attention*, Cambridge, MA: MIT Press.
Vygotsky, L.S. (1962) *Thought and Language*, Cambridge, MA: MIT Press.
Washburn, M.F. (1928) Emotion and thought: A motor theory of their relation, in Murchison, C. (ed.) *Feelings and Emotions: The Wittenberg Symposium*, pp. 99–145, Worcester, MA: Clark University Press.
Watson, J.B. (1924) *Behaviorism*, New York: W.W. Norton.
Webb, T.W. & Graziano M.S.A. (2015) The attention schema theory: A mechanistic account of subjective awareness, *Frontiers in Psychology*, **10**, art. 500.
Webb, T.W., Kean, H.H. & Graziano, M.S.A. (2016) Effects of awareness on the control of attention, *Journal of Cognitive Neuroscience*, **28**, pp. 842–851.
Wolpert, D.M., Ghahramani, Z. & Jordan, M.I. (1995) An internal model for sensorimotor integration, *Science*, **269**, pp. 1880–1882.
Wundt, W. (1900) *Die Sprache*, Leipzig: Engelmann.