Decentering the Brain

Embodied Cognition and the Critique of Neurocentrism and Narrow-Minded Philosophy of Mind

Shaun Gallagher
University of Memphis, USA, and University of Wollongong, Australia • s.gallagher/at/memphis.edu

> Context • Challenges by embodied, enactive, extended and ecological approaches to cognition have provided good reasons to shift away from neurocentric theories. **> Problem** • Classic cognitivist accounts tend towards internalism, representationalism and methodological individualism. Such accounts not only picture the brain as the central and almost exclusive mechanism of cognition, they also conceive of brain function in terms that ignore the dynamical relations among brain, body and environment. **> Method** • I review four areas of research (perception, action/agency, self, social cognition) where enactivist accounts have shown alternative ways of thinking about the brain. **> Results** • Taken together, such analyses form a comprehensive alternative to the classic conceptions of cognitivist, computational neuroscience. **> Implications** • Such considerations motivate the need to re-think our understanding of how the brain itself works. They suggest that the best explanation of brain function may be found in the mixed vocabularies of embodied and situated cognition, developmental psychology, ecological psychology, dynamic systems theory, applied linguistics, the theory of affordances and material engagement, rather than the narrow vocabulary of computational neuroscience. **> Constructivist content** • This account is consistent with an enactivist-constructivist approach to cognition. **> Key words** • **Internalism, perception, agency, autonomy, self, social cognition, predictive processing, enactivism.**

Introduction: Internalism and methodological individualism in cognitive science

« 1 » According to classic cognitivist, i.e., computationalist/internalist theories, the brain is the most central part of the central nervous system, where everything of importance concerning cognition, action, self-awareness and our relations with others happens. Michael Gazzaniga (1998), in his prediction that psychology will be a thing of the past, to be replaced by neuroscience, provides a nice example of an idea that operates in numerous theories of cognition, namely that psychological processes are reducible to neurological processes:

> “My view of how the brain works is rooted in an evolutionary perspective that moves from the fact that our mental life reflects the actions of many, perhaps dozens to thousands, of neural devices that are built into our brains at the factory. These devices do crucial things for us, from managing our walking and breathing to helping us with syllogisms.” (Gazzaniga 1998: xiii)

« 2 » The related claim by Alvin Goldman and Frederique de Vignemont (2009: 154) that the “central system of the mind […i.e.,] the brain is the seat of most, if not all, mental events” represents a similar view. According to this classic orthodoxy, the brain is where emotions happen; it is where intentions are formed, and actions are prepared; it is where our understanding of others takes shape. If there is a self it is in the brain (e.g., Northoff et al. 2006), or as Francis Crick once put it: “You're nothing but a pack of neurons” (Crick 1994: 1). This neurocentric view has become so pervasive it is taught to parents who want to understand their children. Thus, a website[1] devoted to children's health states: “The brain is like a central computer that controls all bodily functions […]” Likewise, according to many researchers, improving education involves learning how the brain learns, since that is where learning happens (e.g., Blakemore & Frith 2005; Jones 2009).

« 3 » In thinking about cognition, self, agency, free will, autonomy, social cognition, and other aspects of mind – and in thinking about when things go wrong in these domains (as in psychopathology) – neurocentrism (or neuroessentialism) is standard. Indeed, recently developed neuro-based disciplines – neurophilosophy, neurotheology, neuroeconomics, neuro-marketing, neuroaesthetics, neuropolitics, neurolaw, neuroeducation, and so on – purportedly help us to explain any topic. Across a number of fields, neurocentrism has strong advocates who defend it as the only or best way to think about such things (e.g., Huber & Kutschenko 2009; Titley, Brunel & Hansel 2017; Shelley 2013; Lee, VanderPloeg & Strifler 2016).

1 | http://kidshealth.org/en/parents/brain-nervous-system.html accessed 17 May 2018.

« 4 » In philosophy, neurocentrism as a form of internalism is represented by the well-known brain-in-the-vat thought experiment, which, beyond its use as a thought experiment, is sometimes proposed as a model for our best understanding of how things work.

> “Some form of internalism must be right because there isn't anything else to do the job. The brain is all we have for the purpose of representing the world to ourselves and everything we can use must be inside the brain. Each of our beliefs must be possible for a being who is a brain in a vat because each of us is precisely a brain in a vat; the vat is a skull and the 'messages' coming in are coming in by way of impacts on the nervous system.” (Searle 1983: 230)

« 5 » In theories of social cognition, we find a complementary focus on methodological individualism – the idea that we can find a complete explanation of how we understand others in a set of mechanisms that are entirely contained in the individual brain – a theory of mind module (ToMM) or a mirror neuron system that automatically simulates the other person's mental states.

« 6 » Neurocentrism can be described as a "narrow" perspective on cognition. The term "narrow" is a technical term in philosophy of mind. It refers to processes contained "in the head" – for example, brain-based representational processes and contents. Narrow-minded views have been challenged by "wide" "E-approaches" – that is, embodied, embedded, extended, enactive, ecological approaches to cognition, which in various ways argue that the unit of explanation ought to be brain-body-environment. On such externalist views, the brain is not dismissed as unimportant for understanding cognition, action, emotion, human experience, and so on; rather it is decentered and given a partial, although still important, role to play along with bodily and environmental factors.

« 7 » In this target article I review four areas or issues where challenges by these E-approaches have provided good reasons to shift away from neurocentric theories: perception, agency and free will, self, and social cognition. Each of these is a complex and multifaceted topic, and I will not be able to do justice to any one of them here.[2] My intent is to paint a large picture and to suggest that, taken together, such analyses form a comprehensive alternative to the classic conceptions of cognitivist, computational neuroscience. If one thinks of these different areas of research as four distinct topics, then one might also think that these apparently dissociated investigations actually converge on some basic assumptions about how to best make sense of cognitive phenomena. I will try to show, as we go along, however, that the convergence of principles reflects deeper connections among these areas. Before turning to these topics, I will discuss the idea that we need to re-think our understanding of how the brain itself works, specifically from the perspective of phylogeny, the importance of which was already intimated by Gazzaniga.

Brain, body and beyond

« 8 » The neural reuse hypothesis is an important and influential insight into how we understand brain functions. As Michael Anderson (2010) explains it, neural circuits originally established for one use can be reused or redeployed for other purposes while still maintaining their original function. This hypothesis was originally understood in terms of an evolutionary notion of plasticity, exaptation: "the shift in the course of evolution of a given trait or mechanism, which is later on reused to serve new purposes and functions" (Gallese 2014: 6). A good example is Broca's area in the human brain. The homologous area in the monkey involves motor functions. Across evolutionary changes it retains these original functions – movement preparation, action sequencing, and action imitation (Binkofski & Buccino 2004). But, in the human, this area is exapted for additional functions involving language and action recognition functions. Its function in speech production has been long known and well established (Broca 1861). The presence of mirror neurons in this area links it to action recognition (Rizzolatti et al. 1996). Mirror neurons are another example of reuse: originally motor neurons involved in motor control, they are exapted in the course of evolution to serve social cognition.

« 9 » This idea of reuse has been appropriated by a narrow, internalist "weak" conception of embodied cognition (EC). According to weak EC, neither the physical body itself (its anatomy, activity, postural body-schematic processes) nor the environment is an important contributory to cognition. Rather, what is important for weak EC are B(ody)-formatted representations and the reuse hypothesis. B-formatted representations are non-propositional interoceptive or motoric representations "of one's own bodily states and activities" (Goldman 2012: 74). These B-formats are characterized as "sanitized" neural representations (Goldman & de Vignemont 2009), and are sometimes discussed under the heading of the "body in the brain" (e.g., Berlucchi & Aglioti 2010). On the weak EC view, the reuse hypothesis is put to use as follows: Any cognitive task that employs a B-formatted representation in either its original function or its exapted/derived function is, on this definition, a form of embodied cognition. Examples include not only mirror neurons and their role in social-cognitive simulation, but also Friedemann Pulvermüller's (2005) language-grounding hypothesis – the idea that action words, like *lick*, *pick*, and *kick* activate cortical motor areas that involve tongue, hand, and foot, respectively. In this case, motor areas and interoceptive, B-formatted motor representations are reused for language processing. Along this same line, by simulation or metaphor, one can explain the embodied roots of abstract thought (Barsalou 2008; Lakoff & Johnson 1999). Thus, "higher-order thought is grounded in low-level representations of motor actions" (Goldman 2014: 94).

« 10 » The evolutionary principle of reuse can get reframed as a developmental principle, in, for example, Stanislas Dehaene's (2005) "neuronal recycling" hypothesis, according to which there are ontogenetic changes in the "visual word form area" of visual cortex when a person learns

2 | These are all themes that I have discussed extensively in other publications. On perception, see Gallagher (2015a, 2017; Gallagher & Zahavi 2012; Hutto et al. in press), on agency, autonomy and free will, see Gallagher (2006, 2007, 2012, 2013a); on self, see Gallagher (2000, 2013b; Gallagher & Daly 2018); and on social cognition, see Gallagher (2001, 2005, 2008a, 2008b).

to read. Goldman also uses the concept of reuse to apply to token neural activations and cognitive events. For example, he mentions "reusing or redeploying B-formats to execute a fundamentally non-bodily cognitive task" (Goldman 2012: 83), an example of which is the activation of mirror neurons, which is "a redeployment of the motoric format in a novel, cognitively interpersonal, task" (ibid: 79).

« 11 » In Gallagher (2015b, 2018), I argued that accepting these extensions in the use of the reuse hypothesis, including Dehaene's (2005) neuronal recycling hypothesis, and Pulvermüller's (2005) language-grounding hypothesis, actually undermines a purely internalist account of cognition – across all timescales: evolutionary, developmental, and the timescale of everyday action.

« 12 » Indeed, accepting the concepts of neuronal reuse implies a *strong* view of EC. On the timescale of evolution, reuse has everything to do with the body – including its morphological features, which are dismissed as trivial by weak EC (Goldman & de Vignemont 2009). Specifically, and obviously, the human brain evolves with the human body. Evolutionary changes in the body that allow for the upright posture, leading to a restructuring of the skull and jaw, allow for a larger brain and for the development of speech. These changes are accompanied by many other morphological changes involving hands, feet, etc., all of which drive evolutionary changes in the brain, and promote reuse.

« 13 » Not only the body, but also physical, social, and cultural environments are important factors, both evolutionarily and developmentally, for any understanding of neural reuse or neuronal recycling. The importance of these non-neural factors is supported by naturalistic research in recent biology seeking to understand "niche construction," "coevolution of culture and genes," or, more generally, the "social brain hypothesis" (Andler 2016: 303–313). Neither brain evolution nor brain development happens *in vitro* or in a vat. The role of the cultural environment, for example, is directly relevant in developmental contexts. This remains unstated, but implicit even in Goldman's discussion of Pulvermüller's work in neural linguistics – "an excellent example of the redeployment of an older (motoric) system, featuring a bodily format, to help execute tasks of language comprehension."[3] Activation of perceptual or motor areas for language and conceptual processing is not just the result of brain plasticity, but cultural practices and learning. The roles of culture and context (including bodily practices and environmental factors), apply equally to token events. For example, motor simulations related to word processing in the context of a sentence are more specific than the meaning represented by the abstract verb outside of a sentence (Naumann 2016): one would expect a different pattern of neural activation for the sentences "Bill picked up the needle" versus "Bill picked up the barbell" – since there are differences in both neural and bodily activations for the differences in grasping (the picking up) involved in such actions. Importantly, the neural activation will depend not only on knowing what a barbell is, or what a needle (or what kind of needle – sewing, compass, hypodermic) is, but also to some significant extent on the history of one's use of such items, and one's skill level, one's bodily practices (consider novice *versus* expert seamstress or weightlifter).

« 14 » Such things are not just neuronal, but also involve bodily, social and cultural factors. This requires a reinterpretation of the notion of reuse. The plasticity involved here is not just neuronal plasticity, but, more importantly, *metaplasticity* (Malafouris 2013); not just brains, but bodies and environments, and social and cultural practices undergo interrelated reuse or plastic changes due to their on-going, dynamical interactions across all relevant timescales.

"More than just evolving (in the restricted Darwinian sense of variation under natural selection), we have been altering our own developmental paths [including our own brains] by making and changing the material means by which we engage the world […]. The plasticity of the mind is embedded and inextricably enfolded with the plasticity of culture." (Malafouris 2015: 351)

3 | Goldman (2014: 103). Note, however, that there have been some replication problems re: semantic somatotopy (Bedny & Caramazza 2011; Willems et al. 2009; Postle et al. 2008; Naumann 2011).

« 15 » Full consideration of the reuse hypothesis leads us directly to the role of body and environment, including cultural context. The unit of explanation is not just the brain, not just the body, not just the environment, but the brain-body-environment (Gallagher et al. 2013). In evolutionary terms, the brain operates the way it does because it is part of an organism that has hands that can reach and grasp in specific ways, and eyes structured to focus, an upright posture, an autonomic system, and so forth, all of which evolved to cope with specific kinds of environments, and with other people. Changes to any component of the individual's bodily, environmental, or experienced social-cultural context will elicit responses from the system as a whole. As the enactivists have argued, rather than internal mental representations or the computation of information, we should understand the brain as participating in the overall action of the system as a whole (Anderson & Chemero 2017; Di Paolo & De Jaegher 2012; Fuchs 2018; Gallagher 2017; Hutto & Myin 2013; Thompson 2007).

« 16 » We can see this broader enactivist notion of reuse in the following four issues. Within each analysis the idea that we can explain the phenomenon mainly in terms of brain processes breaks down and we are led to see the irreducible role of non-neural processes. Moreover, perception, agency/action, self and social cognition are topics thought to be basic and central to most other processes of interest to cognitive science. Understanding how the brain functions within the wider system with respect to these issues will go a long way towards laying the groundwork for a more comprehensive and less neurocentric cognitive science.

Perception

« 17 » On neurocentric theories, perception is something that happens in the brain. Depending on how rigidly one wants to distinguish perception from cognition, on the one side, and action, on the other, perception may be narrowed down to activation of the primary or early perceptual areas (e.g., visual, auditory, tactile cortexes) with meaningful content being added by higher-level cognitive or conceptual representations, and all of this followed by motoric processing leading to action. This is what Susan Hurley (1998) critically called the "sandwich model" of perception, where sensory input leads to higher-order cognition resulting in action output. These processes may be conceived as being more or less dynamically integrated. Perception may be thought to be an inferential process as Hermann von Helmholtz (1867) and more recent predictive models (e.g., Hohwy 2013) would have it, where perception is constituted in top-down predictive processes (informed by generative models on what is typically considered the cognitive level), and, in some cases, active inference (when action is involved).

« 18 » For Jacob Hohwy (2016) and many others who champion Bayesian predictive coding, all such processing is brain-bound, tightly wrapped in a Markov blanket that strictly isolates the brain from body and world. Prediction-error minimization (PEM) in the brain does the important work; active inference (moving around the world) simply serves the central processes:

> "PEM should make us resist conceptions of [a mind-world] relation on which the mind is in some fundamental way open or porous to the world, or on which it is in some strong sense embodied, extended or enactive. Instead, PEM reveals the mind to be inferentially secluded from the world, it seems to be more neurocentrically skull-bound than embodied or extended, and action itself is more an inferential process on sensory input than an enactive coupling with the environment." (Hohwy 2016: 259)

« 19 » In contrast to such internalist conceptions, phenomenology and enactive and extended EC approaches argue that perception should be understood as a set of dynamical processes that relate brain, body, and environment. Inspired by Merleau-Ponty, for example, enactivists argue that the body is involved in at least two ways in perception:

- Perception is closely tied to action, and thus partially constituted by sensorimotor contingencies (Di Paolo, Buhrmann & Barandiaran 2017; O'Regan & Noë 2001; Noë 2004), and
- Perception is shaped by bodily-affective processes (Colombetti 2014; Gallagher 2017; Gallagher & Bower 2014; Thompson & Stapleton 2009).

In the case of sensorimotor contingencies, not only does motor control depend on perceptual input, any movement of one's body changes one's perception. Informed by detailed sensorimotor contingencies, enactive perception is often described in terms of affordances. In the phenomenological philosophers this idea can be traced back to Edmund Husserl's (1989) notion of the "I can." The idea is that I perceive the world in the pragmatic terms of what *I can* do, or in terms of my skill or my expertise. Perception is, as James Gibson (1977) argued, affordance-based. An expert trained in architecture may perceive more affordances than the novice, or different ones; a city-dweller may see the surrounding city environment or a rural pasture differently from how a farmer would.

« 20 » Perception is shaped not only by pragmatic affordances related to sensory-motor contingencies – the "I can" – but also by affective factors of embodiment. Even if I am skilled and capable of grabbing an object, I may not feel "up to the task," or I may not feel motivated or interested. I may not see the object in precisely the same way as I would if I were interested, or if I were not so tired. There may be an affective cost that diminishes what an object affords. Not only does one have a practical (sensorimotor) apprehension of accessibility, but one also has an affective take on that same accessibility, in terms of interest or inclination to follow through or in terms of the ease or difficulty of acting.

« 21 » This affective dimension can be cashed out in terms of a more liberal predictive processing view where the priors or generative models are not reduced to brain-bound processes but can include embodied, affective components. Lisa Barrett and Moshe Bar (also see Barrett & Simmons 2015; Chanes & Barrett 2016), for example, have proposed the "affective prediction hypothesis," which

> "implies that responses signaling an object's salience, relevance or value do not occur as a separate step after the object is identified. Instead, affective responses support vision from the very moment that visual stimulation begins." (Barrett & Bar 2009: 1325)

« 22 » This is not just a matter of internal processing in the brain, but includes a dynamical relation between brain and body. Along with processing in the early visual area, for example, activation of the medial orbital frontal cortex (OFC) initiates a train of muscular and hormonal changes throughout the body. This generates interoceptive feedback from organs, muscles, and joints associated with prior experience, which immediately integrates with current exteroceptive sensory input. This means that the organism as a whole is responding and contributing to perception.

> "The OFC's ongoing integration of sensory information from the external world with that from the body indicates that conscious percepts are indeed intrinsically infused with affective value, so that the affective salience or significance of an object is not computed after the fact. […T]he predictions generated during object perception carry affective value as a necessary and normal part of visual experience. (Barrett & Bar 2009: 1328).

« 23 » Perception involves whole-body dispositions and adjustments and what Patrick Freund et al. (2018) call "anatomically informed priors" (see also Allen & Friston 2018). This implies, first, that perception is not just action-oriented, or recognition-oriented; it is also reward-oriented, hedonic, aesthetic, and affective in the broadest sense. Second, it means that perceptual networks are dynamically connected to and deeply affected by embodied processes that involve multiple systems, such as endocrine and autonomic systems (Gallagher & Allen 2018). Fatigue and hunger, for example, involve extra-neural processes that influence brain function and have an effect on perception. Homeostatic regulation depends on chemical influences in the endocrine system. In hypoglycemic conditions (which can slow or weaken brain function) perception is modulated by complex chemical processes in the body-brain system as it couples with the environment.

« 24 » These considerations lead away from narrow, internalist conceptions of the mind, even in predictive processing (PP) theories. Andy Clark points in this direction:

> “PP thus provides, or so I will argue, the perfect neuro-computational partner for recent work on the embodied mind – work that stresses the constant engagement of the world by cycles of perceptual-motor activity.” (Clark 2016: 1)

In contrast to Hohwy, who makes the “neurocentrically skull-bound” generative model do most of the work, Clark (ibid: 133) emphasizes active inference–active, embodied engagement that manipulates the environment in order to reduce prediction errors. Enactivist interpretations of the predictive model move even more in that direction (Allen & Friston 2018; Bitbol & Gallagher 2018; Bruineberg, Kiverstein & Rietveld 2018; Gallagher & Allen 2018; Ramstead, Badcock & Friston 2018). As Karl Friston puts it:

> “We must here understand ‘model’ in the most inclusive sense, as combining interpretive dispositions, morphology, and neural architecture, and as implying a highly tuned ‘fit’ between the active, embodied organism and the embedded environment.” (Friston et al. 2012: 6)

« 25 » As Friston (2013: 213) summarizes, “an agent does not have a model of its world – it is a model.” With this it is not clear that we still need to think of the brain as requiring its own model of the world, or even that we need to keep the concept of a *model*. Rather than a generative model, which implies an additional internal dynamics separate from bodily and environmental processes, we can refer to a generative dynamics coordinated across brain, body and environment. To develop a conception of enactive perception, we need to understand active inference in terms of *action* rather than prediction error minimization (see Bruineberg, Kiverstein & Rietveld 2018).

Action, agency, and autonomy

« 26 » The well-known experiments by Benjamin Libet (1985, 1992) suggested that if free will does exist it is to be found in processes that span the 150 milliseconds of neuronal activation occurring just prior to issue of a motor command. Libet showed that we become conscious of the decision or urge to move only after some 500–850 milliseconds of brain activity (the “readiness potential,” which correlates with preparation for that specific action) have already occurred. This suggests that consciousness does not play a role in causing the action, at least until approximately 150 milliseconds before motor activation. We should note two things about this result. First, this has nothing to do with free will. I have argued that Libet’s experiments were about motor control processes that typically remain unconscious, and that free will involves larger timescales and factors that cannot be reduced to neuronal processes. I will return to this point. Second, 150 milliseconds of neuronal activation in pre-motor processes is plenty of time to generate a sense of agency that is experienced, pre-reflectively, as intrinsic to one’s action. This has been a standard way to think about the sense of agency (Haggard & Magno 1999; Haggard & Eimer 1999; Gallagher 2000) – as something that anticipates the action itself generated in neuronal processes that just precede motor command.

« 27 » The sense of agency, however, is more complex than just this experience of motor control. It also includes a sense of what one is doing in the world – a sense of what one’s action is accomplishing (Gallagher 2012; Haggard 2017). Even more than this, one’s experience of agency depends on a variety of factors that go beyond physical bodily action or its immediate intentional aspect. It may include the scope of affordances available to the agent in specific environments; it can also include prior intention formation that may benefit from communications with other people, as well as retrospective attribution that may take shape in narrative. Action is always situated in physical, social and cultural circumstances. Importantly, other people and social forces have an effect on one’s sense of agency. Even on the pre-reflective level, the presence of others can have an effect on my perception of action possibilities. An agent may be both capable of and proficient at performing action A, for example, throwing a basketball through a hoop. Nonetheless her performance, and her sense of agency, may be negatively affected simply by the fact that there is an audience of basketball superstars watching her. She may in fact feel a degree of inadequacy in such circumstances, simply because a specific set of people are present. Likewise, and in contrast to many analyses of agency in philosophy of mind and the cognitive sciences, deliberation, intention formation and motivation to act are not simply mental states in one’s head, or causal brain states. Rather, they are often processes or states co-constituted with others in processes of communication. Consider also the effects of peer pressure, implicit or explicit social referencing, or one’s habitual behavior in the presence of others. Such phenomena may detract from or increase one’s feeling of agency and ability to act. It is also the case that specific types of long-standing social arrangements, such as apartheid, can have prolonged effects on a person’s (or a people’s) long-term sense of agency, essentially robbing them of possibilities for action (Gallagher 2012).

« 28 » Returning to the question of free will, I have argued that we should not think of the exercise of free will as equivalent to the initiation and control of bodily movement (mental causation), which is the target of the Libet experiments, and the standard way of thinking of free will from Descartes to many contemporary philosophers and neuroscientists (Gallagher 2006). Motor control, the body-schematic details of which we are not usually conscious, is not the same thing as the exercise of free will. The consciousness that pertains to action is not (as in the Libet experiments) focused on deciding to move one’s body. Rather, awareness of bodily movement is typically minimal and recessive.

« 29 » Given the prevalence of the traditional conception of free will (as involving mental causation of bodily movement), it may be productive to shift to the concept of autonomy. The notion of autonomy, at least, is not associated with abstract motor processes that make up intentional actions; it applies to an engaged, situated agent and intentional actions themselves, described at an appropriate pragmatic level of description. Immanuel Kant (1996) is the *locus classicus* for the traditional conception of autonomy, which involves self-sufficiency, self-legislation, or self-determination. Following this tradition, most contemporary discussions of autonomy take it to be an individualistic concept. Harry Frankfurt (1982), for example, frames it in terms of rational-reflective decision-making processes – a deliberation leading to the formation of second-order intentions or desires. Other theorists take narrative competency to be an important part of the precise kind of self-reflection that informs decision-making. David Velleman (2005), for example, argues that narrative-based reflection provides a framework for forming and testing one's intentions and for guiding actions and the formation of self-identity. This allows for autonomous self-governance, which depends on forming intentions that are consistent with one's narrative understanding of oneself.

« 30 » Embodied action, however, happens in a world that is not only physical but also social. Our actions and our decisions often involve other people; they are often joint actions steered by physical and social forces and affordances. In this respect, autonomy is relational (Christman 2004; Mackenzie & Stoljar 2000). In contrast to traditional models of an autonomous individual, Catriona Mackenzie and Natalie Stoljar (2000: 4) suggest that we think of autonomy "as a characteristic of agents who are emotional, embodied, desiring, creative, and feeling, as well as rational, creatures." Indeed, our deliberations, and our intentions may be formulated in communicative practices, and may not be reducible to processes that are exclusive to one individual. In that sense they may be co-constituted in one's interactions with others. They may be shaped by institutional forces, social practices, and normative expectations. Accordingly, autonomy is a matter of degree, something that is enhanced or reduced by various physical, social, economic, cultural factors, our relations with others, as well as our own narrative practices. Individuals are always embedded in social contexts, characterized by intersubjective and normative relations that can either enhance or impoverish the control they have over their lives and can expand or constrict their action possibilities. For this reason, it is best to conceive of agency, intention, and autonomy in these embodied and socially situated terms, rather than in terms of brain processes measured in milliseconds.

« 31 » Autonomy, as I mentioned, is closely connected to notions of self-determination, self-legislation, and self-sufficiency, and even if these are understood in relational terms, it refers us directly to some concept of *self*. To understand action and agency we need to understand the agent, and precisely what it is that constitutes the agent.

Self

« 32 » There is a long tradition of looking for the self in the brain. Even those who would deny that the self is reducible to brain processes consider the self to have a special relation to the brain. Karl Popper and John Eccles (1977), for example, defend a dualism that takes the self to be an autonomous entity that significantly interacts with and controls neural processes. "The self-conscious mind acts upon […] neural centres, modifying the dynamic spatio-temporal patterns of the neural events" (Popper & Eccles 1977: 495).

« 33 » There is still great interest in how various aspects of self relate to brain, or how specific brain areas correlate with self-related phenomena. Self-referential processes, including autobiographical knowledge, personal beliefs, self-conceptions, and face self-recognition are related to left hemisphere activity (Turk et al. 2003; see also Kircher et al. 2000) or right frontal cortex (Platek et al. 2003), or right lateral parietal cortex (Lou et al. 2004) or medial prefrontal cortex bilaterally (Fossati et al. 2003). Moreover, cortical midline structures (CMS) process information related to self when subjects reflectively think about themselves, or when they make judgments about their own personalities (D'Argembeau et al. 2007; Gutchess et al. 2007; Northoff & Bermpohl 2004; Northoff et al. 2006; Ruby et al. 2009). Northoff contends that the CMS includes a unitary neural network responsible for all self-related phenomena (Northoff et al. 2006). The CMS also connects to subcortical areas, suggesting a relation to an embodied self (Northoff & Panksepp 2008).

« 34 » Given the diversity and large number of cortical areas correlated with self-reference, Seth Gillihan and Martha Farah (2005) were led to suggest that there is no specialized brain area responsible for generating "the self." Dorothée Legrand and Perrine Ruby (2009) argue in a complementary way that no area of the brain is exclusively self-specific since "every significant activation in the [self condition] was also found in either the [other person condition] or the [general semantic] condition, or both" (Craik et al. 1999: 30; also Gillihan & Farah 2005: 94). It thus seems right to say that the self is both everywhere and nowhere in the brain (Vogeley & Gallagher 2011). It is not just that the brain is so complex, however, but also that the concept of self is ambiguous. Accordingly, in any analysis of self we need to define the precise aspect of the self under study. Selves consist of a variety of aspects – experiential, ecological, agentive – and are capable of various forms of self-recognition, self-related cognition, self-narrative, and self-specific perception and action. In this respect, selves are more "in-the-world" than "in-the-brain" (Vogeley & Gallagher 2011: 129).

« 35 » In contrast to theories that would reduce the self to one particular type of thing – for example, a self-model generated by neuronal processes, and nothing more (e.g., Metzinger 2004); or the abstract product of narratives, and nothing more (e.g., Dennett 1991); or nothing more than a 3-second-long experience *sans* body or agency or narrative (Strawson 1999) – pluralist theories suggest that the self is many things. William James (1950), for example, distinguished between physical, social, and private selves. Ulrich Neisser (1988) distinguished ecological, interpersonal, conceptual, extended, and private selves. In an attempt to capture the plurality of self-related factors and the idea that the agentive self is more "in-the-world" than "in the brain," I have proposed a pattern theory of self (Gallagher 2013b). In brief, the pattern theory of self (PTS) argues that a self is constituted as a pattern or dynamical *Gestalt* comprised of a sufficient number of characteristic factors, including embodied, experiential, affective, behavioral, intersubjective, psychological/cognitive, reflective, narrative, extended and normative factors (see Table 1). It is important to note that this is not an additive list of factors; rather these components or aspects are dynamically interrelated in a pattern or *Gestalt* arrangement (Gallagher & Daly 2018). Accordingly, a change in one element, above a certain threshold, will lead, via dynamical interactions, to changes in others. For example, as suggested above, aspects of self-experience, such as the sense of

Elements of the pattern	Brief description
Embodied elements	Core biological, ecological and interoceptive factors, allowing the system to distinguish between itself and what is not itself – extremely basic to all kinds of animal behavior.
Minimal experiential elements	First-person, pre-reflective, conscious experience, reflecting the self/non-self distinction, manifest in various sensory-motor modalities (kinaesthesia, proprioception, touch, vision, etc.) – including a *sense of ownership* (the "mineness" of one's experience) and a *sense of agency* for one's actions (Gallagher 2012; Rochat 2011).
Affective aspects	Affect/emotion/temperament, ranging from bodily affects to what may be a typical affective or emotion pattern (Newen, Welpinghus & Juckel 2015).
Behavioral aspects	Behaviors and actions make us who we are – behavioral habits reflect, and perhaps actually constitute, our character. This is a classic view that goes back at least to Aristotle.
Intersubjective interactions and capacities	Human are born with a capacity for attuning to intersubjective existence, which develops into a social self-consciousness – a self-for-others (Mead 1913), manifested behaviorally in mirror self-recognition (Gallup, Anderson & Platek 2011), joint actions and communicative practices.
Psychological/cognitive elements	Traditional theories of the self focus on these factors, which may range from explicit self-consciousness to a conceptual understanding of self as self, to personality traits of which one may not be self-conscious at all – psychological continuity and the importance of memory are highlighted in the literature on personal identity.
Reflective capacities	The ability to reflect on one's experiences and actions – closely related to the notions of autonomy and moral personhood, including the capacity to reflect and form second-order volitions about one's desires (Frankfurt 1982; Taylor 1989).
Narrative capacities	Although some theorists make the strong claim that narratives are constitutive for selves (Schechtman 2011), for PTS one can lose the ability to construct a self-narrative (as in cases of dysnarrativa) and still remain a self to the extent that other elements of the pattern remain in place.
Extended/situated elements	Including the possibilities presented by physical pieces of property, and various things that we own (James 1950). Not only may we identify with our material belongings, or the technologies we use, our professions and the institutions we work in, but we are also dynamically related to the action possibilities they afford.
Normative factors	Ranging across possibilities presented by the kind of family structure and situation in which we grew up to cultural and normative practices, involving physical and mental health, gender, race, and economic status, that define our way of living.

Table 1 • Dynamical aspects of the self-pattern (from Gallagher & Daly 2018).

agency, can be modulated by other complex, relational aspects, such as social, normative factors that involve culture, gender, race, health, etc., and by specific intersubjective factors that can either diminish or enhance one's autonomy and sense of agency. There is much more to say about PTS and about how to investigate the dynamical aspects of the self-pattern (see Gallagher & Daly 2018). Here, however, my intention is simply to note that a self-pattern is more than a neural pattern. This does not mean that there is no connection between self and brain. Indeed, changes in neurophysiology can "index" changes in the self-pattern (Fingelkurts & Fingelkurts 2017).

« 36 » Within the framework of predictive processing (PP), Jakub Limanowski and Felix Blankenburg (2013: 1), for example, argue that the minimal (pre-reflective) experiential aspects of the bodily self can be "mapped onto a hierarchical generative model […] and may constitute the basis for higher-level, cognitive forms of self-referral." On the same model affective factors may involve multisensory integration (Seth 2013) that also relate to self-recognition (Apps & Tsakiris 2014). Matthew Apps and Manos Tsakiris also note the influence of culturally shaped priors on PP:

> "There is also evidence of more long-term contextual influences on self-recognition related priors, highlighted by the role that cultural and societal effects have on self-other decision-making. For instance, self-other face recognition has been shown to be different across cultures […]." (ibid: 14)

Once again, these dynamical neural and extra-neural integrations reflect the various dynamical relations between the embodied and experiential aspects of the self-pattern and the extended and normative aspects. To paraphrase Friston, the agent does not have a self, it *is* a self, where the self is not a model in the brain but a pattern of generative dynamics coordinated across the elements of brain, body and environment.

« 37 » Importantly, on a PP approach, self-specific neural processing may arise in any multisensory processing, thereby avoiding problems (outlined above) involved in positing specialized circuits or parts of the brain that are self-specific. Indeed, Apps and

Tsakiris (2014: 8) claim that a PP account "provides flexibility, with fewer constraints on what types of information can drive self-recognition." The strong claim is that predictive models can explain all of the various factors that contribute to the self-pattern. As they note,

> "This is particularly important, given the evidence to suggest that the continuity of the self may be underpinned by many different types of information, the integration of which leads to a coherent sense of one's body." (ibid: 9)

« 38 » To be clear, this type of analysis sends us back to issues previously discussed (in the section on perception) about how we might best interpret predictive processing models. Although Hohwy and John Michael (2017) build an internalist PP model of self, it is interesting to note that they see their account as consistent with PP accounts of minimal phenomenal (experiential) selfhood, or self model (Limanowski & Blankenburg 2013; Metzinger 2004), bodily self-awareness (Apps & Tsakiris 2014), interoceptive aspects (Seth, Suzuki & Critchley 2011), intersubjective aspects that relate self and other (Moutoussis et al. 2014), social understanding of self and other (Frith & Friston 2015), and psychodynamical notions of self (e.g., Fotopoulou 2012), thereby touching on many elements of the self-pattern. All of this, however, on their account, is reducible to the brain's predictive model. Rather than taking the fully embodied self to be the agent of active inference, or part of a dynamical system that includes the brain, Hohwy and Michael take the body to be a representation in the internal model of the agent:

> "The body is nothing special, it is just one among many causes interacting with each other in the environment, and in the course of this impacting on the senses. Representation of the body is nothing special either; it is just one among many causes that get represented in the internal model used for prediction error minimization." (Hohwy & Michael 2017: 367f)

« 39 » In contrast to reducing the self to neuronal patterns, or to the patterns of inference that constitute a self-model, PTS argues that the self-pattern is a "real pattern" (Dennett 1991) of dynamical relations among brain-body-environment (broadly speaking). Such dynamical relations may be partially indexed or traced by neuronal processes to the extent that the latter partially underpin various factors of the self-pattern, but they are not reducible to such processes. One important component of this pattern includes the agent's intersubjective interactions and capacities for social cognition, phenomena that are clearly more than just brain processes.

Social cognition

« 40 » Standard approaches to social cognition (theory of mind – ToM) typically accept the assumption of methodological individualism. Although ToM seeks to explain how we understand the minds of others, via mindreading processes, all those processes are said to be contained in the observer's own head/brain. Either a theory of mind module (ToMM) in pre-frontal areas activates subpersonal processes that constitute a "tacit theory" or implicit use of folk psychology (Carruthers 2015), or mirror neurons generate subpersonal processes that simulate the actions and minds of others (Gallese 2001; Goldman 2006).

« 41 » The alternative phenomenological-enactive approach to social cognition, interaction theory (IT), rejects the idea that we mindread the hidden mental states of others. It focuses on embodied interaction processes that draw on multiple semiotic resources in the other's postures, movements, gestures, facial expressions, vocal intonations, communicative practices and actions in contextualized social and cultural environments. This includes an enactive model of direct social perception of the other person's embodied mental states (especially intentions and emotions) (De Jaegher, Di Paolo & Gallagher 2010; Gallagher 2005, 2008a; Reddy 2008). For example, if emotional episodes, as Giovanna Colombetti suggests:

> "correspond to specific self-organizing forms or second-order constraints – *emotion forms*, as I call them – that recruit or entrain various processes (neural, muscular, autonomic, etc.) into highly integrated configurations or patterns." (Colombetti 2014: 69)

and if these patterns include bodily expressions, comportments, and actions, then (consistent with what we said in previous sections) my perception of another person's emotions may be considered a form of perceptual pattern recognition – one that is action- (or interaction-) oriented and takes that pattern as a social affordance for further response on my part. On this view we are not engaged in third-person observation of others, but in second-person ("I-thou" or "we-mode") interactions. In contrast to methodological individualism, what does the work of social cognition are not mechanisms internal to the individual, but our engaged interactions that happen in our shared, intersubjective world and that build on shared or reciprocal social affordances.

« 42 » Scientific evidence for this is found in developmental studies of infants, from birth onward, in their face-to-face, primary intersubjective relations, and their pragmatically contextualized secondary intersubjective relations (Trevarthen 1979). Primary intersubjectivity involves innate or early-developing sensory-motor capacities that bring us into relation with others and allow us to interact with them. In part, these capacities involve action and perceptual experience – we are able to *see* or more generally *perceive* in the other person's bodily postures, movements, gestures, facial expressions, gaze direction, vocal intonation, etc. what they intend and what they feel. We respond with our own bodily movements, gestures, facial expressions, gaze, etc. On this view, the other's mind is in her embodied comportment, and manifests itself in second-person interactions. For infants, these highly embodied and situated interactions form the basis for a developing understanding of others.

« 43 » Infants already have a sense, from their own self-movement and proprioception, of their own agency, and they see this kind of agency in others. They respond, interactively, to certain kinds of entities (specifically to other agents) in the environment. They can respond in a distinctive way to human faces (Johnson 2000; Johnson, Slaughter & Carey 1998; Legerstee 1991), for example. From birth infants are capable of perceiving and responding to facial gestures presented by others, and seem to be directly attuned to the actions and gestures of other humans (Meltzoff & Moore 1977, 1994; Gallagher & Meltzoff 1996). Although claims about neonatal imitation remain controversial, this is not a worry for IT. Whether it is differential imitation or a mere arousal response (Anisfeld 2005; Keven & Akins 2017; Jones 2006, 2009; but see Nagy et al. 2013; Nagy, Pal & Orvos 2014; Vincini et al. 2017a, 2017b; Vincini & Jhang 2018) it nonetheless leads infant and caregiver to intersubjective interaction. An initial adult facial gesture may motivate the infant's arousal and response; in turn the infant's response has an effect on the adult who is encouraged to continue with facial games, etc. In this way, even a mere arousal response could facilitate early social interaction.

« 44 » Primary intersubjectivity can be specified in much more detail. At 2 months, infants are already attuned to the other person's attention; they follow the other's head movements and gaze (Baron-Cohen 1995; Maurer & Barrera 1981). Also at 2 months, second-person *interaction* is evidenced by the timing and emotional response of infants' behavior (Gopnik & Meltzoff 1997). This is part of a mutual attunement that characterizes interactions and that can be specified in detail in their dynamical relations and the integration of the intrinsic temporalities of the agents' movements (Trevarthen 1999; Trevarthen et al. 2006). At 5–7 months infants can detect visual-audio correspondences specifying the expression of emotions (Walker 1982; Hobson 1993, 2002). At 6 months they see grasping as goal-directed. At 10–11 months infants can parse intentional boundaries within some kinds of continuous action (Baldwin & Baird 2001; Baird & Baldwin 2001; Woodward & Sommerville 2000).

« 45 » Such expressions, intonations, gestures, and movements do not float freely in the air; they are situated in the world, anchored to specific contexts. Accordingly, towards the end of the first year, infants start to notice how others engage with the world. For joint attention and secondary intersubjectivity context becomes very important, and it helps us to intersubjectively co-constitute the meaning of the world.

« 46 » A good example of secondary intersubjectivity can be found in conversation analysis, the rich analysis of speech acts situated in circumstances that involve our own and others' postures and movements, along with environmental arrangements and affordances (Goodwin 2000, 2017). Charles Goodwin shows that meaning emerges in action and interaction, specifically at the intersection of social, cultural, material and temporal structures of the environment. Meaning is accomplished, not just via speech but by drawing on "different kinds of semiotic resources" available in the environment and in whole-body pragmatics. "For example, spoken language builds signs within the stream of speech, gestures use the body in a particular way, while posture and orientation use the body in another, etc." Goodwin emphasizes the "visible, public deployment of multiple semiotic fields that mutually elaborate each other" (ibid: 1494): vocal intonations (some of which have a deontic rather than descriptive force); movements, postures and bodily orientations; instituted norms; references to completed actions; interruption of activities, and so on.

« 47 » As an example, Goodwin provides a detailed analysis of a dispute between two young girls over a game of hopscotch.

> “Unlike talk, gestures can't be heard. [This means] Carla [one of the girls] actively works to position her hand gestures so that they will be perceived by Diana [the other girl] […] Carla's hand is explicitly positioned in Diana's line of sight […] thrusting the gesturing hand toward Diana's face twists Carla's body into a configuration in which her hand, arm and the upper part of her torso are actually leaning toward Diana.” (Goodwin 2001: 1498)

« 48 » The proximity of the gesture to the other girl's face has meaning. If it were a touch rather than a gesture, how hard or soft, and where the touch occurred would also have meaning.

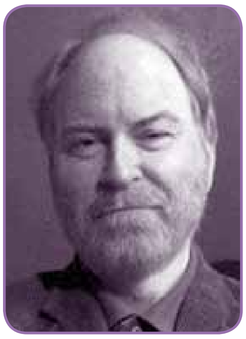
« 49 » Importantly, interaction is not one-sided. A response draws a further response. In the interaction, the conversation is not confined to vocalization and gesture – reference is made to the surrounding environment, joint attention is established and then broken and then re-established. The accomplishment of meaning is not under the control of just one individual; rather it depends upon two-way interaction. According to IT, social understanding builds on precisely this complex integration of primary and secondary intersubjective capacities, situated within pragmatic and social contexts, supplemented with and supporting communicative and narrative processes. In this regard, it is the interaction itself that contributes something not reducible to the actions of the individuals involved, or to individual brain states (De Jaegher, Di Paolo & Gallagher 2010).

« 50 » IT does not deny that the brain is an important part of the body or that it plays an important role in cognition and social cognition. Indeed, mirror neurons, motor control processes, and notions of reuse may play some role in explaining social cognition. It is reasonable to think that reuse is in some way constrained by original use. If so, then the fact that our perceptual-motor systems were originally designed for *action*, rather than for observation, is significant (Anderson & Chemero 2017). It is likely that this action orientation carries through to the reuse of our motor systems in contexts of social cognition, but again (as I indicated in §§13–14) this requires a reinterpretation of reuse in relation to wider contexts involving metaplasticity. Thus, I see your action as an affordance that motivates my own action – I see it as something I can respond to in broader contexts of social interactions, joint actions, cultural practices, etc., and that is precisely how I understand your action.

« 51 » The interaction that is essential to social cognition is not reducible to the interaction of neurons; it requires agentive bodies, and others, situated in physical, social, cultural and normatively constrained environments in support of interactions that happen in the world rather than in individual brains.

Conclusion

« 52 » Since the 1990s, the assumption in cognitive science has been that neuroscience will at some point replace psychology and that we will adjust our philosophies of mind accordingly (Gazzaniga 1998). The expectation was that the best explanation of brain function would be worked out in the vocabulary of neuroscience. In contrast, I want to suggest that the best explanation of brain function may be found in the vocabularies of embodied and situated cognition,



SHAUN GALLAGHER
is the Lillian and Morrie Moss Professor of Excellence in Philosophy at the University of Memphis. His areas of research include phenomenology and the cognitive sciences, especially topics related to embodiment, self, agency and intersubjectivity. Gallagher holds a secondary research appointment at the University of Wollongong, Australia, and is Honorary Professor of Health Science at the University of Tromsø, Norway. He has held visiting positions at Cambridge; Copenhagen; the Ecole Normale Supériure, Paris and Lyon; the Humboldt University in Berlin, and most recently at Keble College, University of Oxford. Professor Gallagher received the Humboldt Foundation's Anneliese Maier Research Award (2012–18). His publications include *How the Body Shapes the Mind* (2005), *The Phenomenological Mind* (with Dan Zahavi, 2008, second edition 2012); and *Enactivist Interventions: Rethinking the Mind* (2017).

developmental psychology, ecological psychology, dynamic systems theory, applied linguistics, the theory of affordances, along with the anthropological insights found in discussions that extend from concepts of cultural niche to material engagement. There is a methodological question involved here: whether neuroscience can start to speak this different language and enter into the right kind of dialogue. There is also a substantial question: how do brains operate in the complex and dynamical mix of interactions that involve perceiving, moving, gesturing, acting, emoting and expressing bodies?

> “It is conventional to think of the nervous system as an organ that monitors and motivates the body rather than an organ *controlled by the body* […]. Nevertheless, the body's influence on the nervous system is as important for the organism as is neural dominion over the body. (Purves 1988: 1)

« 53 » That the body essentially constrains and “pushes” the organization of the brain through its dynamic behavioral interaction with the environment was already well documented by George Coghill (1929), in *Anatomy and the Problem of Behavior*. Maurice Merleau-Ponty, in his 1957–58 lectures on the concept of nature, was inspired by Coghill's work for setting the principles of dynamic anatomy in opposition to strict determinism, and he provides an appropriate conclusion.

> “The nervous system emerges from a preneural dynamic. Thus when the nervous excitation occurs, it can't play an important role in the organization of the nervous system. This organization is not so much due to the functioning of the neuron as to the growth of the total organism. The preneural system of integration ‘strides across’ the nervous functioning and it doesn't stop when it appears. So the nervous system can't be the ultimate explanation. Then we must admit an intrinsic potentiality of growth, a dynamic system reacting to its surroundings as an organism would do. It replaces the function of conduction as being a consequence, not a principle of the system.” (Merleau-Ponty 2003: 192)

« 54 » In this article I have considered a sampling of research areas in the cognitive sciences – perception, agency, self and social cognition. In each case I have argued against a narrow or neurocentric reductionism. These are basic phenomena upon which many cognitive capacities are built. One could easily see the same principles at work in a number of other areas where more comprehensive accounts have been developed by taking embodied, ecological, enactive and extended approaches seriously – for example, research on memory (Sutton 2010), expert performance (Høffding 2015; Ilundáin-Agurruza 2016), collective intentionality (Tollefsen & Gallagher 2017), psychopathology (Gallagher 2013c), and psychotherapy (Garcia & Di Paolo 2018; Röhricht et al. 2014). In each case one can acknowledge the importance of what the brain is doing, operating as part of a larger circuit that includes body and environment. The brain is not at the center of a circle with radii of control extending to other elements; it is one component arranged in the circuit, or in what Viktor von Weizsäcker (1986) called a *Gestalt* circle of brain, body, and the (physical, social, cultural) environment.

Acknowledgements

Research for this article was supported by the Humboldt Foundation's Anneliese Maier Research Award and the Australian Research Council (ARC) grant, *Minds in Skilled Performance*. DP170102987. I benefitted greatly from the comments of two blind, but insightful reviewers.

References

Allen M. & Friston K. J. (2018) From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. Synthese 195(6): 2459–2482. ▸ http://cepa.info/4099

Anderson M. L. (2010) Neural reuse: A fundamental reorganizing principle of the brain. Behavioral and Brain Sciences 33: 245–266.

Anderson M. & Chemero A. (2017) The brain evolved to guide action. In: Shepherd S. V. (ed.) The Wiley handbook of evolutionary neuroscience John Wiley, New York: 1–20.

Andler D. (2016) La silhouette de l'humain. Quelle place pour le naturalisme dans le monde d'aujourd'hui? Editions Gallimard, Paris.

Anisfeld M. (2005) No compelling evidence to dispute Piaget's timetable of the development of representational imitation in infancy. In: Rizzolatti G., Hurley S. & Chater N. (eds.) Perspectives on imitation: From neuroscience to social science. MIT Press, Cambridge MA: 107–131.

Apps M. A. & Tsakiris M. (2014) The free-energy self: A predictive coding account of self-recognition. Neuroscience & Biobehavioral Reviews 41: 85–97.

Baird J. A. & Baldwin D. A. (2001) Making sense of human behavior: Action parsing and intentional inference. In: Malle B. F., Moses L. J. & Baldwin D. A. (eds.) Intentions and intentionality: Foundations of social cognition. MIT Press, Cambridge MA.

Baldwin D. A. & Baird J. A. (2001) Discerning intentions in dynamic human action. Trends in Cognitive Science 5(4): 171–78.

Baron-Cohen S. (1995) Mindblindness: An essay on autism and theory of mind. MIT Press, Cambridge MA.

Barrett L. F. & Bar M. (2009) See it with feeling: Affective predictions during object perception. Philosophical Transactions of the Royal Society B 364(1521): 1325–1334.

Barrett L. F. & Simmons W. K. (2015) Interoceptive predictions in the brain. Nature Reviews Neuroscience 16(7): 419–429.

Barsalou L. W. (2008) Grounding cognition. Annual Review of Psychology 59: 617–645.

Bedny M. & Caramazza A. (2011) Perception, action and word meanings in the human brain: The case from action verbs. Annals of the New York Academy of Sciences 1224: 81–95.

Berlucchi G. & Aglioti S. M. (2010) The body in the brain revisited. Experimental brain research 200(1): 25.

Binkofski F. & Buccino G. (2004) Motor functions of the Broca's region. Brain and language 89(2): 362–369.

Bitbol M. & Gallagher S. (2018) Autopoiesis and the free energy principle. Comment on Ramsted, Badcock and Friston. Physics of Life Review 24: 24–26. ▸ http://cepa.info/5407

Blakemore S.-J. & Frith U. (2005) The learning brain: Lessons for education. Blackwell Publishing, Malden.

Broca P. (1861) Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech) Bulletin de la Société Anatomique 6: 330–357.

Bruineberg J., Kiverstein J. & Rietveld E. (2018) The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. Synthese 195(6): 2417–2444. ▸ http://cepa.info/4497

Carruthers P. (2015) Perceiving mental states. Consciousness and Cognition 36: 498–507.

Chanes L. & Barrett L. F. (2016) Redefining the role of limbic areas in cortical processing. Trends in Cognitive Sciences 20(2): 96–106.

Christman J. (2004) Relational autonomy, liberal individualism, and the social constitution of selves. Philosophical Studies 117(1): 143–164.

Clark A. (2016) Surfing uncertainty: Prediction, action and the embodied mind. Oxford University Press, Oxford.

Coghill G. E. (1929) Anatomy and the problem of behaviour. Cambridge University Press, Cambridge MA.

Colombetti G. (2014) The feeling body: Affective science meets the enactive mind. MIT Press, Cambridge MA.

Craik F. I. M., Moroz T. M., Moscovitch M., Stuss D. T., Winocur G., Tulving E. & Kapur S. (1999) In search of the self: A positron emission tomography study. Psychological Science 10: 26–34.

Crick F. (1994) The astonishing hypothesis: The scientific search for the soul. Charles Scribner's Sons, New York.

De Jaegher H., Di Paolo E. & Gallagher S. (2010) Does social interaction constitute social cognition? Trends in Cognitive Sciences 14(10): 441–447. ▸ http://cepa.info/4349

Dehaene S. (2005) Evolution of human cortical circuits for reading and arithmetic: The "neuronal recycling" hypothesis. In: Dehaene S., Duhamel J.-R., Hauser M. D. & Rizzolatti G. (eds.) From monkey brain to human brain. MIT Press, Cambridge MA: 131–157.

Dennett D. C. (1991) Consciousness explained. Little, Brown & Co, New York.

Di Paolo E., Buhrmann T. & Barandiaran X. (2017) Sensorimotor life: An enactive proposal. Oxford University Press, Oxford.

Di Paolo E. A. & De Jaegher H. (2012) The interactive brain hypothesis. Frontiers in human neuroscience 6:163. ▸ http://cepa.info/761

D'Argembeau A., Ruby P., Collette F., Degueldre C., Balteau E., Luxen A., Maquet P. & Salmon E. (2007) Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective-taking. Journal of Cognitive Neuroscience 19: 935–944.

Fingelkurts A. A. & Fingelkurts A. A. (2017) Three-dimensional components of selfhood in treatment-naive patients with major depressive disorder: A resting-state qEEG imaging study. Neuropsychologia 99: 30–36.

Fossati P., Hevenor S. J., Graham S. J., Grady C., Keightley M. L., Craik F. & Mayberg H. (2003) In search of the emotional self: An fMRI study using positive and negative emotional words. American Journal of Psychiatry 160: 1938–1945.

Fotopoulou A. (2012) Towards psychodynamic neuroscience. In: Fotopoulou A., Conway M. & Pfaff D. (eds.) From the couch to the lab: Trends in psychodynamic neuroscience. Oxford University Press, New York: 25–47.

Frankfurt H. (1982) The importance of what we care about. Synthese 53(2): 257–272.

Freund P., Friston K., Thompson A. J., Stephan K. E., Ashburner J., Bach D. R., Nagy Z., Helms G., Draganski B., Mohammadi S., Schwab M. E., Curt A. & Weiskopf N. (2018) Embodied neurology: An integrative framework for neurological disorders. Brain 139(6): 1855–1861.

Friston K., Adams R. A., Perrinet L. & Breakspear M. (2012) Perceptions as hypotheses: Saccades as experiments. Frontiers in Psychology 3: 151.

Frith C. & Friston K. (2015) A duet for one. Consciousness and Cognition 36: 390–405.

Fuchs T. (2018) Ecology of the brain: The phenomenology and biology of the embodied mind. Oxford: Oxford University Press.

Gallagher S. (2000) Philosophical conceptions of the self: Implications for cognitive science. Trends in Cognitive Sciences 4(1): 14–21. ▸ http://cepa.info/4360

Gallagher S. (2001) The practice of mind: Theory, simulation, or primary interaction? Journal of Consciousness Studies 8(5–7): 83–107.

Gallagher S. (2005) How the body shapes the mind. Oxford University Press, Oxford.

Gallagher S. (2006) Where's the action? Epiphenomenalism and the problem of free will. In: Banks W., Pockett S. & Gallagher S. (eds.) Does consciousness cause behavior? An

investigation of the nature of volition. MIT Press, Cambridge MA: 109–124.
Gallagher S. (2007) The natural philosophy of agency. Philosophy Compass 2(2): 347–357.
Gallagher S. (2008a) Direct perception in the intersubjective context. Consciousness and Cognition 17: 535–543.
Gallagher S. (2008b) Inference or interaction: Social cognition without precursors. Philosophical Explorations 11(3): 163–173.
Gallagher S. (2012) Multiple aspects of agency. New Ideas in Psychology 30: 15–31.
Gallagher S. (2013a) Ambiguity in the sense of agency. In: Clark A., Kiverstein J. & Vierkant T. (eds.) Decomposing the will. Oxford University Press, Oxford: 118–135.
Gallagher S. (2013b) A pattern theory of self. Frontiers in Human Neuroscience 7(443): 1–7.
Gallagher S. (2013c) Intersubjectivity and psychopathology. In: Fulford B., Davies M., Graham G., Sadler J. & Stanghellini G. (eds.) Oxford Handbook of philosophy of psychiatry. Oxford University Press, Oxford: 258–274.
Gallagher S. (2015a) Seeing things in the right way: Normativity in social perception. In: Doyon M. & Breyer T. (eds.) Normativity in Perception. Palgrave-Macmillan, Basingstoke: 117–127.
Gallagher S. (2015b) Reuse and body-formatted representations in simulation theory. Cognitive Systems Research 34–35: 35–43.
Gallagher S. (2017) Enactivist interventions: Rethinking the mind. Oxford University Press, Oxford.
Gallagher S. (2018) Building a stronger concept of embodiment. In: Newen A., de Bruin L. & Gallagher S. (eds.) The Oxford handbook of 4E cognition. Oxford University Press, Oxford: 353–367.
Gallagher S. & Allen M. (2018) Active inference, enactivism and the hermeneutics of social cognition. Synthese 195(6): 2627–2648. ▸ http://cepa.info/4222
Gallagher S. & Bower M. (2014) Making enactivism even more embodied. AVANT 5(2): 232–247. ▸ http://cepa.info/4495
Gallagher S. & Daly A. (2018) Dynamical relations in the self-pattern. Frontiers in Psychology 9: 664.
Gallagher S., Hutto D., Slaby J. & Cole J. (2013) The brain as part of an enactive system. Behavioral and Brain Sciences 36(4): 421–422. ▸ http://cepa.info/2510
Gallagher S. & Meltzoff A. (1996) The earliest sense of self and others: Merleau-Ponty and recent developmental studies. Philosophical Psychology 9(2): 213–236.
Gallagher S. & Zahavi D. (2012) The phenomenological mind. Routledge, London.
Gallese V. (2001) The "shared manifold" hypothesis: From mirror neurons to empathy. Journal of Consciousness Studies 8: 33–50.
Gallese V. (2014) Bodily selves in relation: Embodied simulation as second-person perspective on intersubjectivity. Philosophical Transactions of the Royal Society B: Biological Sciences 369: 20130177.
Gallup G. G., Anderson J. R. & Platek S. M. (2011) Self-recognition. In: Gallagher S. (ed.) The Oxford handbook of the self. Oxford University Press, Oxford: 80–110.
Garcia E. & Di Paolo E. (2018) Embodied coordination and psychotherapeutic outcome: Beyond simple mappings. Frontiers in Psychology 9: 1257. https://doi.org/10.3389/fpsyg.2018.01257.
Gazzaniga M. S. (1998) The mind's past. University of California Press, Berkeley.
Gibson J. J. (1977) The theory of affordances. In: Shaw R. & Bransford J. (eds.) Perceiving, acting, and knowing. Erlbaum, Hillsdale NJ: 67–82.
Gillihan S. J. & Farah M. J. (2005) Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. Psychological Bulletin 131(1): 76–97.
Goldman A. (2006) Simulating minds: The philosophy, psychology and neuroscience of mindreading. Oxford University Press, Oxford.
Goldman A. I. (2012) A moderate approach to embodied cognitive science. Review of Philosophy and Psychology 3(1): 71–88.
Goldman A. I. (2014) The bodily formats approach to embodied cognition. In: Kriegel U. (ed.) Current controversies in philosophy of mind. Routledge, New York: 91–108. ▸ http://cepa.info/5191
Goldman A. & de Vignemont F. (2009) Is social cognition embodied? Trends in Cognitive Sciences 13(4): 154–159.
Goodwin C. (2000) Action and embodiment within situated human interaction. Journal of Pragmatics 32: 1489–1522.
Goodwin C. (2017) Co-operative action. Cambridge University Press, Cambridge.
Gopnik A. & Meltzoff A. N. (1997) Words, thoughts and theories. MIT Press, Cambridge MA.
Gutchess A. H., Kensinger E. A., Yoon C., Schacter D. L. (2007) Ageing and the self-reference effect in memory. Memory 15: 822–837.
Haggard P. (2017) Sense of agency in the human brain. Nature Reviews Neuroscience 18(4): 196–207.
Haggard P. & Eimer M. (1999) On the relation between brain potentials and the awareness of voluntary movements. Experimental Brain Research 126: 128–133.
Haggard P. & Magno E. (1999) Localising awareness of action with transcranial magnetic stimulation. Experimental Brain Research 127: 102–107.
Helmholtz H. von (1867) Handbuch der physiologischen Optik. Leopold Voss, Leipzig. English translation: Helmholtz H. (2005) Treatise on physiological optics. Edited by J. Southall. Dover Publications, Mineola NY.
Hobson P. (1993) The emotional origins of social understanding. Philosophical Psychology 6: 227–249.
Hobson P. (2002) The cradle of thought. Macmillan, London.
Hohwy J. (2013) The predictive mind. Oxford University Press, Oxford.
Hohwy J. (2016) The self-evidencing brain. Noûs 50(2): 259–285.
Hohwy J. & Michael J. (2017) Why should any body have a self. In: De Vignemont F. & Alsmith A. (eds.) The subject's matter. MIT Press, Cambridge MA: 363–391.
Huber L. & Kutschenko L. K. (2009) Medicine in a neurocentric world: About the explanatory power of neuroscientific models in medical research and practice. Medicine Studies 1(4): 307–313.
Hurley S. L. (1998) Consciousness in action. Harvard University Press, Cambridge MA.
Husserl E. (1989) Ideas pertaining to a pure phenomenology and to a phenomenological philosophy – Second Book: Studies in the phenomenology of constitution. Translated by Rojcewicz R. & Schuwer A. Kluwer Academic, Dordrecht.
Hutto D., Gallagher S., Ilundáin-Agurruza J. & Hipólito I. (in press) Culture in mind – An enactivist account: Not cognitive penetration but cultural permeation. In: Kirmayer L. J., Kitayama S., Worthman C. M., Lemelson R. & Cummings C. A. (eds.) Culture, mind,

and brain: Emerging concepts, models, applications. Cambridge University Press, New York NY.

Hutto D. D. & Myin E. (2013) Radicalizing enactivism: Basic minds without content. MIT Press, Cambridge MA.

Høffding S. (2015) A phenomenology of expert musicianship. PhD thesis, Department of Philosophy, University of Copenhagen, Copenhagen.

Ilundáin-Agurruza J. (2016) Holism and the cultivation of excellence in sports and performance. Routledge, London.

James W. (1950) Principles of Psychology. Dover, New York. Originally published in 1890.

Johnson S. C. (2000) The recognition of mentalistic agents in infancy. Trends in Cognitive Science 4(1): 22–28.

Johnson S., Slaughter V. & Carey S. (1998) Whose gaze will infants follow? The elicitation of gaze-following in 12-month-old infants. Developmental Science 1(2): 233–238.

Jones S. S. (2006) Exploration or imitation? The effect of music on 4-week-old infants' tongue protrusions. Infant Behavior and Development 29(1): 126–130.

Jones P. H. (2009) Introducing neuroeducational research: Neuroscience, education and the brain from contexts to practice. Routledge, London.

Jones S. S. (2009) The development of imitation in infancy. Philosophical Transactions of the Royal Society of London B: Biological Sciences 364(1528): 2325–2335.

Kant I. (1996) Practical philosophy. Edited and translated by M. J. Gregor. Cambridge University Press, Cambridge.

Keven N. & Akins K. A. (2017) Neonatal imitation in context: Sensory-motor development in the perinatal period. Behavioral and Brain Sciences 40: E381.

Kircher T. T. J., Senior C., Phillips M. L., Benson P. J., Bullmore E. T., Brammer M., Simmons A., Williams S. C. R., Bartels M. & David A. S. (2000) Towards a functional neuroanatomy of self processing: Effects of faces and words. Cognitive Brain Research 10: 133–144.

Lakoff G. & Johnson M. (1999) Philosophy in the flesh: The embodied mind and its challenge to western thought. Basic Books, New York.

Lee S. E., VanderPloeg A. & Strifler J. (2016) The neurocentric view on consumers of cotton: An application of consumer neuroscience. In: Proceedings of the ITAA Annual Conference. Annual Conference Proceedings #73. Iowa State University Digital Repository: 1–2. https://lib.dr.iastate.edu/itaa_proceedings/2016/presentations/98.

Legerstee M. (1991) The role of person and object in eliciting early imitation. Journal of Experimental Child Psychology 51(3): 423–33.

Legrand D. & Ruby P. (2009) What is self specific? A theoretical investigation and a critical review of neuroimaging results. Psychological Review 116(1): 252–282.

Libet B. (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. Behavioral and Brain Sciences 8: 529–566.

Libet B. (1992) The neural time-factor in perception, volition, and free will. Revue de Métaphysique et de Morale 2: 255–272.

Limanowski J. & Blankenburg F. (2013) Minimal self-models and the free energy principle. Frontiers in human neuroscience 7: 547.

Lou H. C., Luber B., Crupain M., Keenan J. P., Nowak M., Kjaer T. W., Sackeim H. A. & Lisanby S. H. (2004) Parietal cortex and representation of the mental self. Proceedings of the National Academy of Sciences USA 101: 6827–6832.

Mackenzie C. & Stoljar N. (2000) Introduction: Autonomy refigured. In: Mackenzie C. & Stoljar N. (eds.) Relational autonomy: Feminist perspectives on autonomy, agency, and the social self. Oxford University Press, Oxford: 3–31.

Malafouris L. (2013) How things shape the mind. MIT Press, Cambridge MA.

Malafouris L. (2015) Metaplasticity and the primacy of material engagement. Time and Mind 8(4): 351–371.

Maurer D. & Barrera M. E. (1981) Infants' perception of natural and distorted arrangements of a schematic face. Child Development 52(1): 196–202.

Mead G. H. (1913) The social self. Journal of Philosophy, Psychology, and Scientific Methods 10(14): 374–380.

Meltzoff A. N. & Moore M. K. (1977) Imitation of facial and manual gestures by human neonates. Science 198(4312): 75–78.

Meltzoff A. & Moore M. K. (1994) Imitation, memory, and the representation of persons, Infant Behavior and Development 17: 83–99.

Merleau-Ponty M. (2003) Nature: Course notes from the College de France, ed. R. Vallier. Northwestern University Press, Evanston.

Metzinger T. (2004) Being no one: The self-model theory of subjectivity. MIT Press, Cambridge.

Moutoussis M., Fearon P., El-Deredy W., Dolan R. J. & Friston K. J. (2014) Bayesian inferences about the self (and others): A review. Consciousness and Cognition 25: 67–76.

Nagy E., Pal A. & Orvos H. (2014) Learning to imitate individual finger movements by the human neonate. Developmental Science 17: 841–57.

Nagy E., Pilling K., Orvos H. & Molnar P. (2013) Imitation of tongue protrusion in human neonates: Specificity of the response in a large sample. Developmental Psychology 49(9): 1628.

Naumann R. (2016) Dynamics in the brain and dynamic frame theory for action verbs. In: Brunel L., Versace R. & Brouillet D. (eds.) The sensory nature of knowledge. Proceedings of the SMCLC-11. Düsseldorf University Press, Düsseldorf: 109–130.

Neisser U. (1988) Five kinds of self-knowledge. Philosophical Psychology 1: 35–59.

Newen A., Welpinghus A. & Juckel G. (2015) Emotion recognition as pattern recognition: The relevance of perception. Mind & Language 30(2): 187–208.

Northoff G. & Bermpohl F. (2004) Cortical midline structures and the self. Trends in Cognitive Sciences 8: 102–107.

Northoff G., Heinzel A., de Greck M., Bennpohl F., Dobrowolny H. & Panksepp J. (2006) Self-referential processing in our brain: A meta-analysis of imaging studies on the self. Neuroimage 31: 440–457.

Northoff G. & Panksepp J. (2008) The trans-species concept of self and the subcortical–cortical midline system. Trends in Cognitive Sciences 12(7): 259–264.

Noë A. (2004) Action in perception. MIT Press, Cambridge MA.

O'Regan J. K. & Noë A. (2001) A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24(5): 939–973. ▸ http://cepa.info/2285

Platek S. M., Myers T. E., Critton S. R. & Gallup G. G. (2003) A left-hand advantage for self-description: The impact of schizotypal personality traits. Schizophrenia Research 65: 147–151.

Popper K. R. & Eccles J. C. (1977) The self and its brain: An argument for interactionism. Springer International, New York.

Postle N., McMahon K. L., Ashton R., Meredith M. & de Zubicaray G. I. (2008) Action word meaning representations in cyto-architectonically defined primary and premotor cortex. Neuroimage 43: 634–644.

Pulvermüller F. (2005) Brain mechanisms linking language and action. Nature Reviews Neuroscience 6: 576–582.

Purves D. (1988) Body and brain: A trophic theory of neural connections. Harvard University Press, Cambridge MA.

Ramstead M. J. D., Badcock P. B. & Friston K. J. (2018) Answering Schrödinger's question: A free-energy formulation. Physics of Life Review 24: 1–16.

Reddy V. (2008) How infants know minds. Harvard University Press, Cambridge MA.

Rizzolatti G., Fadiga L., Gallese V. & Fogassi L. (1996) Premotor cortex and the recognition of motor actions. Cognitive Brain Research 3(2): 131–141.

Rochat P. (2011) What is it like to be a newborn. In: Gallagher S. (ed.) The Oxford handbook of the self. Oxford University Press, Oxford: 57–79.

Ruby P., Collette F., D'Argembeau A., Peters F., Degueldre C., Balteau E., Luxen A., Maquet P., Salmon E. (2009) Perspective taking to assess self-personality: What's modified in Alzheimer's disease? Neurobiology of Aging 30: 1637–1651.

Röhricht F., Gallagher S., Geuter U. & Hutto D. D. (2014) Embodied cognition and body psychotherapy: The construction of new therapeutic environments. Sensoria: A Journal of Mind, Brain & Culture 10(1): 11–20.

Schechtman M. (2011) The narrative self. In: Gallagher S. (ed.) The Oxford handbook of the self. Oxford University Press, Oxford: 394–416.

Searle J. R. (1983) Intentionality: An essay in the philosophy of mind. Cambridge University Press, Cambridge.

Seth A. K. (2013) Interoceptive inference, emotion, and the embodied self. Trends in Cognitive Sciences 17(11): 565–573. ► http://cepa.info/4518

Seth A. K., Suzuki K. & Critchley H. D. (2011) An interoceptive predictive coding model of conscious presence. Frontiers in Psychology 2: 395.

Shelley B. P. (2013) Brain wellbeing and brain-healthy lifestyle in the neurocentric age. Archives of Medicine and Heath Sciences 1(2): 99–102.

Strawson G. (1999) The self. In: Gallagher S. & Shear J. (eds.) Models of the self. Imprint Academic, Exeter: 1–24.

Sutton J. (2010) Exograms and interdisciplinarity: History, the extended mind and the civilizing process. In: Menary R. (ed.) The extended mind. MIT Press, Cambridge MA: 189–225.

Taylor C. (1989) Sources of the self: The making of the modern identity. Harvard University Press, Cambridge MA.

Thompson E. (2007) Mind in life: Biology, phenomenology, and the sciences of mind. Harvard University Press, Cambridge MA. ► http://constructivist.info/3/2/117

Thompson E. & Stapleton M. (2009) Making sense of sense-making: Reflections on enactive and extended mind theories. Topoi 28(1): 23–30. ► http://cepa.info/2290

Titley H. K., Brunel N. & Hansel C. (2017) Toward a neurocentric view of learning. Neuron 95(1): 19–32.

Tollefsen D. & Gallagher S. (2017) We-narratives and the stability and depth of shared agency. Philosophy of the Social Sciences 47(2): 95–110.

Trevarthen C. (1979) Communication and cooperation in early infancy: A description of primary intersubjectivity. In: Bullowa M. (ed.) Before Speech: The Beginning of Interpersonal Communication. Cambridge University Press, Cambridge.

Trevarthen C. (1999) Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. Musicae scientiae 3(1 suppl): 155–215.

Trevarthen C., Aitken K. J., Nagy E., Delafield-Butt J. T. & Vandekerckhove M. (2006) Collaborative regulations of vitality in early childhood: Stress in intimate relationships and postnatal psychopathology. In: Cicchetti D. & Cohen D. J. (ed.) Developmental Psychopathology. John Wiley and Sons, New York: 65–126.

Turk D. J., Heatherton T. F., Macrae C. N., Kelley W. M. & Gazzaniga M. S. (2003) Out of contact, out of mind: The distributed nature of the self. Annals of the New York Academy of Science 1001: 65–78.

Velleman J. D. (2005) Self as narrator. In: Christman J. & Anderson J. (eds.) Autonomy and the challenges to liberalism: New Essays. Cambridge University Press, New York: 56–76.

Vincini S. & Jhang Y. (2018) Association but not recognition: An alternative model for differential imitation from 0 to 2 months. Review of Philosophy and Psychology 9(2): 395–427.

Vincini S., Jhang Y., Buder E. H. & Gallagher S. (2017a) An unsettled debate: Key empirical and theoretical questions are still open (Commentary on Keven and Akins) Behavioral and Brain Sciences 40: E401.

Vincini S., Jhang Y., Buder E. H. & Gallagher S. (2017b) Neonatal imitation: Theory, experimental design and significance for the field of social cognition. Frontiers in Psychology – Cognitive Science 8: 1323.

Vogeley K. & Gallagher S. (2011) The self in the brain. In: Gallagher S. (ed.) The Oxford handbook of the self. Oxford University Press, Oxford: 111–136.

Walker A. S. (1982) Intermodal perception of expressive behaviors by human infants. Journal of Experimental Child Psychology 33(3): 514–35.

Weizsäcker V. von. (1986) Der Gestaltkreis: Theorie der Einheit von Wahrnehmen und Bewegen. Fifth edition. Thieme, Stuttgart.

Willems R. M., Toni I., Hagoort P. & Casasanto D. (2009) Neural dissociations between action verb understanding and motor imagery. Journal of Cognitive Neuroscience 22: 2387–2400.

Woodward A. L. & Sommerville J. A. (2000) Twelve-month-old infants interpret action in context. Psychological Science 11: 73–77.

Received: 11 June 2018
Accepted: 4 September 2018