# The 'I' of the beholder: studying the 'self' across the humanities and neuroscience

Alessia Pannese

#### Correspondence to

Alessia Pannese, Italian Academy for Advanced Studies in America, Columbia University, 1161 Amsterdam Avenue, New York, NY 10027, USA; ap2215@columbia.edu

Accepted 17 June 2011 Published Online First 23 August 2011

#### ABSTRACT

Long debated within the humanistic tradition, the concept of 'self' has recently been embraced as a subject of investigation by cognitive neuroscience. Tracing parallels between ancient philosophical ideas and current-day scientific research on the 'self', the author proposes that contemporary knowledge based on empirical neuroscientific evidence may inform novel perspectives on—and draw inspiration from—notions grounded in ancient intuitions and traditionally falling within humanistic fields of enquiry. Further, the author suggests that the 'self', as a major object of philosophical and psychological enquiry, as well as a central component of human motivation, cognition, affect and social identity, is an inherently crossdisciplinary research topic, which, by virtue of its pervasive and defining presence in human existence, lends itself—and demands—to be approached both from scientific (objective) and phenomenological (subjective) vantage points. On this premise, the author proposes that the study of the 'self' provides both neuroscientists and humanists with a fertile ground for cross-disciplinary research, and with the challenge and the opportunity to rethink the relationship of science to knowledge.

The notion of 'self' occupies a prominent position in the intellectual debate as a traditional topic for philosophical thought and as a radical and provocative newly adopted object of scientific enquiry. This dual presence in both the humanistic and the scientific domains makes the 'self' a potentially lively and stimulating area of interdisciplinary research. This work is an attempt to highlight this potentiality. By building conceptual bridges between ancient ideas and contemporary knowledge pertaining to the notion of 'self', I will argue that neuroscientific research may both inform and be informed by ideas developed within the humanistic enquiry, and suggest that intellectual openness to rigorous interdisciplinary debate is a desirable scholarly feature, as it provides both neuroscientists and humanists with the challenge and the opportunity to rethink the relationship of their specific approach to knowledge and to (other forms of) knowledge at large.

# THE 'SELF': AN ANCIENT TOPIC FOR CONTEMPORARY RESEARCH

The concept of 'self' has been a topic of interest and a productive line of enquiry since antiquity, in both the East and the West. One of the earliest explicit notions of 'self' is contained in the juxtaposition between *Atman* (individual self) and

Brahman (universal spirit) found in the Upanishads, 1 philosophical treatises composed in India around the first millennium BC. Similar emphasis on the centrality of self-knowledge to human experience emanates from the injunction 'know thyself', which, according to the Greek geographer Pausanias<sup>2</sup> (second century AD), was chiselled into the front porch of Apollo's temple in Delphi, greeting those seeking guidance, and reminding them of the need to first know themselves before they could fit that guidance to their own situations. The theme of self-knowledge (and lack thereof) occupies an equally prominent position in the Socratic thought (ca. 469-399 BC) (as recorded in his disciples' writings<sup>3</sup>), and represents the main source of suffering in Aeschylus' (ca. 525–456 BC)<sup>4</sup> Sophocles'  $(496-406 BC)^5$ (eg, Agamemnon and Oedipus the King).

In ancient times, numerous philosophical theorisations bearing relevance to the debate about the nature of the 'self' have been inscribed in a dualistic framework, whereby the human individual is endowed with two distinct attributes: one material, referred to as 'body' or res extensa, and one immaterial, termed 'soul' (Plato), res cogitans (Descartes), 'mind' or 'spirit' (Hegel). Although, based on the equation of thought with being (I think, therefore I am), the 'self' falls within the domain of the res cogitans, much of the scientific debate surrounding it owes its conceptual origin to an intuition largely pertaining to the res extensa: it was indeed the Cartesian notion of reflex, the automatic bodily response to an external stimulus observed independently of the individual's awareness (such as the knee jerk), which, by extending Galilei's mechanical conception of the physical universe to human behaviour, first conceived the individual as a system responding to external stimuli, thereby laying the conceptual foundation for a science of the mind. Descartes' mentalistic theorisation<sup>6</sup>—postulating the 'self' as a distinct entity, separate from the world, which it is nevertheless in a position to know—is echoed in Locke's concept of consciousness<sup>7</sup> as the perception of what passes in a man's own mind-'[...] consciousness always accompanies thinking, and it is that which makes every one to be what he calls self, and thereby distinguishes himself from all other thinking things' (An Essay Concerning Human Understanding, p. 226)—in Hume's notion of 'self'8 as an illusorily coherent 'bundle of perceptions' (A Treatise of Human Nature, sect. 205, p. 170), as well as in numerous psychological theories developed in the 20th century, whereby the 'self' has been recurrently described as a composite mental construct,  $^{9-14}$  and a complex neuro-socio-cognitive

## Original article

process involving the capacity to become the object of one's own attention, <sup>15</sup> to actively identify, process and store information about oneself, <sup>16</sup> to experience awareness of both private experiences (eg, one's own mental states) and public self-characteristics (eg, one's own body, behaviour, physical appearance, reputation), to maintain a coherent sense of continuous identity across time (autonoetic awareness), to experience agency over one's thoughts and actions, and to be aware of one's existing as an independent and unique entity, distinct from the environment. <sup>17</sup>

Given the heterogeneity of these definitions, in order to illustrate my argument—that the study of the 'self' constitutes a potential terrain for interdisciplinary exchange between neuroscience and the humanities—with specific examples, I will hone the parameters of the debate by focusing on one particular aspect of the problem: the relationship between the sense of 'self' and the face, a recurrent theme in both the scientific and the humanistic literature on the topic.

### THE FACE AS MANIFESTED 'SELF'

The centrality of the face to the debate about selfhood has long been established, dating back millennia. In Greco-roman antiquity it was believed that apprehension of one's inner 'self' occurred by looking at one's face, regarded as the manifestation of the individual's personality. The tight coupling between face and identity is highlighted in the Greek language by the existence of a single term—prosopon—meaning both 'face' and 'mask', thus implying a notion of inextricability between external appearance and characterial traits, and highlighting the importance for humans to develop the ability to 'read' the face as carrier of crucial information about the individual's 'self'. This ancient intuition whereby faces, as embodied carriers of potentially biologically relevant information, command the observer's special attunement, resonates with empirical evidence from neuroscience showing that humans are extremely efficient at recognising faces, an ability they acquire early during childhood. 18 At a neurophysiological level, face recognition has been shown to result from multiple cognitive processes, including an early (100 ms after stimulus onset) categorisation—in which the object is classified as being a face 19 20—followed by a late (170 ms after stimulus onset) identification proper—in which the face is classified as belonging to a specific individual.  $^{19\ 21\ 22}$ This modular account is consistent with neuropsychological  $^{18}$  and functional MRI $^{23}$  evidence of distinct processing for different facial attributes, suggesting that face features are first coded separately, and later integrated into a global representation.  $^{24\ 25\ i}$  Besides temporal aspects, brain specificity for face processing has been found in the spatial domain. Neuroimaging studies have shown that viewing a face is associated with selective engagement of a specific population of neurons in the inferotemporal region of the human brain.<sup>26</sup> In macaques, where neurons selectively responsive to faces are located in both the inferior temporal gyrus and on the banks of the superior temporal sulcus, the former have been shown to be selective for individual identity, while the latter tend to be selective for facial expression,<sup>27</sup> direction of gaze and head orientation.<sup>28</sup> If we

For an alternative account of face recognition based on shared processing resources see Baudouin JY, Tiberghien G. Gender is a dimension of face recognition. *J Exp Psychol Learn Mem Cogn* 2002;28:362—5; Ganel T, Goshen-Gottstein Y. Perceptual integrality of sex and identity of faces: further evidence for the single-route hypothesis. *J Exp Psychol Hum Percept Perform* 2002;28:854—67; Cellerino A, Borghetti D, Valenzano DR, *et al.* Neurophysiological correlates for the perception of facial sexual dimorphism. *Brain Res Bull* 2007;71:515—22.

posit that the fusiform face area reported in humans<sup>26</sup> is the homologue of the inferotemporal region described in macaques, then it is conceivable that the human fusiform face area may be involved in the discrimination of facial identity too.<sup>29</sup> <sup>30</sup>

The overall picture emerging from this combined evidence characterises face recognition as a complex cognitive achievement, entailing multiple specialised computations performed by distinct brain areas. 31-34 This high proficiency in processing faces suggests an evolutionary adaptation, the high biological meaningfulness of the face commanding dedication of increasingly sophisticated processing resources, ultimately resulting in a high degree of neural specialisation for face perception. I propose that this evolutionary account, whereby humans developed an adaptive attunement to reading faces as a result of faces carrying socially and ecologically highly relevant information about other individuals' personality and intentions, fits well with the notion, widespread since Greco-roman antiquity, of face as the embodiment of the 'self'. I suggest that this example eloquently illustrates how neuroscientific evidence gathered in the last two decades may invite novel perspectives on-and provide empirical support to—ancient intuitions dating back over two millennia.

#### THE FACE AS REFLECTED 'SELF'

In the last several decades, researchers have attempted to register the sense of 'self' through mirror self-recognition, adopting an experimental approach pioneered by Charles Darwin, 35 standardised by Gordon Gallup, 36 and later widely applied to both humans and other animals. Early studies belonging to this line of research have shown that humans are normally able to recognise themselves in a mirror by the age of 18–24 months. <sup>37 38</sup> whereas most other animals lack this ability, and placed in front of a mirror behave as if confronted with another animal. 36 39 Also, unlike animals, humans possess a complex language, thought to be a necessary attribute for the development of introspection and self-mentalisation. These considerations are often taken to suggest that only humans are endowed with the intellectual abilities required to experience a sense of 'self', other animals only possessing basic sensory awareness. However, this 'discontinuity theory', postulating a qualitative difference between humans and other animals, clashes with evidence of self-recognition found in non-human primates. Studies have shown that trained chimpanzees,  $^{36}$  orangutans  $^{40}$  as well as pigeons, <sup>41</sup> dolphins <sup>42</sup> and elephants <sup>43</sup> can recognise themselves in a mirror. These findings show that language is not needed to pass the mirror test, and argue in favour of a 'continuity theory', whereby a quantitative (rather than qualitative) difference in the sense of 'self' exists between humans and other animals; self-consciousness being a graded phenomenon whose intensity is correlated with the philogenetic (more evolved species have higher self-consciousness) and ontogenetic (adults are more conscious than newborn) differentiation of the cerebral cortex.44

Mirror studies rely on the assumption that passing the mirror test implies self-recognition and that this, in turn, is a proof of self-consciousness. <sup>39</sup> <sup>45</sup> Although several lines of evidence support the idea that self-recognition may in fact be an epiphenomenon of self-consciousness (eg, the fact that the human timeline for self-recognition correlates with that of the development of deceptive abilities and of emotions requiring self-monitoring from a third-person perspective, such as pride, shame and guilt, all of which imply a clear self—non-self distinction), <sup>38</sup> <sup>46</sup> the validity of the mirror test as a tool to infer

the sense of 'self' has been challenged both theoretically  $^{47\ 48}$  and methodologically,  $^{49-52}$  with critics expressing scepticism towards necessarily coupling self-recognition with any cognitive ability beyond mere kinetic intelligence. According to this argument, all that is needed in order to 'pass' the mirror test is a basic understanding of what a mirror does, and the capacity to match visual information with motor behaviour. Despite generating resonance in the scientific literature, the argument reducing self-recognition to a mere matching of motor cues with visual feedback, with no bearing to a global sense of 'self', is challenged by recent evidence for self-specific<sup>53</sup> or at least selfpreferential (as proposed by Brédart, <sup>54</sup> in consideration of Calder and Young's finding<sup>55</sup>) cross-modal facilitation, which suggests that self-face recognition is the manifestation of a multi-modal phenomenon, thus supporting, and refining through the notion of gradation, Gallup's equation of self-recognition and selfconsciousness.

The study of mirror-self-recognition offers a further example of how the study of the 'self' is conducive to fruitful interaction between scientific and humanistic enquiry. Indeed, Darwin's and Gallup's idea to infer self-knowledge through the use of mirrors is not entirely novel, but rather represents a clever adaptation of an ancient notion. Reflecting devices enabling self-knowledge through apprehension of one's external appearance have been used as a literary construct to address the issue of identity and selfhood since Greco-roman times. The paradigm of reflection can be found in the Greek myth of Narcissus, <sup>56</sup> an exceptionally handsome youth, who, as a divine punishment for spurning those who loved him, was cursed to fall in love with his own face reflected in a pool, and wasted away to death for not being able to leave the beauty of his own reflection (Metamorphoses, Book 3; figure 1). A similarly relevant association to mirrored images is found in the mythological character of Medusa, whose power is initially represented and ultimately diminished through reflecting surfaces (ibid., Book 4). Mirror reflections play an equally pivotal role in Spenser's representation of Britomart,<sup>57</sup> who falls in love with Arthegall upon first seeing his face in her father's magic mirror (Britomart, Book III, II, p. 35). Likewise, the theme of self-knowledge through reflection dominates Milton's literary depiction of Eve, 58 who, shortly after being created, becomes enraptured by an image (her own, unknown to Eve) reflected in water. Interestingly, the voice that makes Eve aware of her own reflected self ('What there thou seest, fair Creature, is thyself') is also the one that promptly urges her away from it ('but follow me, And I will bring thee where no shadow stays'), encouraging her to focus not on her own beauty, but on her male consort ('he Whose image thou art; him thou shall enjoy') (Paradise Lost, Book 4, p. 144 vv. 466-470). Notably, besides highlighting the centrality of the face to the debate about selfhood, Milton's account of God's initially explanatory but soon proscriptive intervention on Eve calls into question the issue of gender in the literary construct of the 'self', as, by being urged to look away from her own reflection, Eve is both denied the pleasure of gazing upon her own beauty—whereas Adam is allowed to delight in it—and forbidden the knowledge of her own identity. It has been argued<sup>59</sup> that because Eve neither recognises nor names herself, her self-knowledge is indirect, as she can know herself only in relation to Adam. I propose that Walker's inferential leap from absence of self-recognition to absence of self-knowledge may be interpreted as closely resonating with Gallup's equation between self-recognition and selfconsciousness, thus exposing conceptual shunts between literary criticism and neuroscience.



**Figure 1** Michelangelo Merisi da Caravaggio (1572—1610). *Narcissus* (1596). Oil on canvas. Roma, Galleria Nazionale d'Arte Antica. The painting depicts the mythological character who, according to the legend, as a divine punishment for disdaining those who loved him, was cursed to fall in love with his reflected image, and wasted away to death for not being able to leave the beauty of his own reflection. In modern neuroscientific terms, Narcissus's tragedy may be reinterpreted as lying not in his vanity and self-love, but in his being deprived of the ability to recognise the reflection of his own face, a condition observed in association with focal brain damage.

#### THE FACE AND THE OBJECTIVE VERSUS SUBJECTIVE 'SELF'

In Greco-roman antiquity, when an introspective and selfmentalising mind was neither conceivable nor experimentally testable, it was believed that apprehension of one's inner 'self' occurred by looking at one's face. This ancient intuition bears striking relevance to the modern scientific paradigm, predicated on the acquisition of knowledge through appraisal of objective, visible signs. Today, this externalist approach—postulating that knowledge may be acquired from 'objective' factors found outside the psychological state of the individual subject who acquires it—is commonly declined in a variety of experimental practices involving behavioural, psychological, psychophysical and neuroimaging methods, most of which seek evidence for the sense of 'self' as evoked through visual exposure to one's own face. The resulting body of work, building on the objectivistic assumption that the sense of 'self' is associated with observable and measurable behavioural and neural correlates—the so-called astonishing hypothesis 60—has highlighted both potentials and pitfalls. Indeed, notwithstanding the abundance of behavioural, 61–63 imaging 64–69 and lesion studies, 70–74 the neural mechanisms involved in the processing of one's own face remain controversial. Furthermore, moving beyond the merely descriptive level, it is unclear how behavioural and neural activity give rise to the subjective experience of 'self'.

Reflecting the unease with the externalist/objectivist stance, the notion that the sense of 'self' is suitable to be investigated scientifically has been radically challenged both on philosophical and on epistemological grounds. Descartes<sup>75</sup> prominently opposed the externalistic view with an internalistic perspective. whereby the 'self' is separate from the external world, which it can only perceive through (fallible) senses, leading to the impossibility of achieving true knowledge of it. Likewise, Hume's 'self'8 is conceived as inextricable from its fluctuating and incessantly evolving perceptual experience. Wittgenstein too, in his lifelong concerns with matters of the mind, <sup>76</sup> regularly touched on the question of the 'self' (indiscriminately referred to as 'selbst', 'Ich', 'Seele', 'Subjekt'), which he early on framed in radically solipsistic terms—'There really is only one world soul, which I for preference call my soul [...]' (Notebooks 1914–1916, p. 49)<sup>77</sup>—that were maintained until his (36-year-) later writings—'But it is still false to say: [...] I is a different person from L.W.' (as reported by H. Sluga, 1996, p. 320).<sup>78</sup> By denying the 'self' the possibility of being an object, Wittgenstein's solipsistic antiobjectivism rejects any science of the 'self', since 'science deals with objects' (ibid., p. 329). A similar stance has more recently been articulated by Berrios and Markova<sup>79</sup> on the ground of the 'self' being a mere construct of Western thought, rather than a natural entity, which makes it intractable through scientific methods. The naturalistic perspective, implying that the 'self' can be investigated with third-person methodologies, has been criticised for missing its subjective phenomenological—component. In his seminal essay 'What is it like to be a bat?', Nagel<sup>80</sup> highlights the limitations of reductionist attempts at explaining phenomenal mental experiences exclusively in terms of physical events taking place in the brain, and emphasises the necessity to also address subjectivity: 'After all, what would be left of what it was like to be a bat if one removed the viewpoint of the bat?'. Zahavi, 81 likewise, argues that due to the subjective nature of the 'self' inferring selfconsciousness through empirical methods such as Gallup's mirror test is intrinsically impossible. Any investigation of the self, he insists, must take the first-person perspective and focus on the experiential givenness of the 'self', a position echoing the main tenet of philosophical phenomenology, in the tradition of Husserl, Heidegger, Sartre and Merleau-Ponty.

Hence, the question remains fundamentally unanswered—or, more accurately, open to a multiplicity of interpretations: is the sense of 'self' an objectively observable physical phenomenon or a scientifically ungraspable subjective mental state? I argue that it is a bit of both, uniquely positioned at the intersection between objective and subjective, scientific and humanistic epistemological approaches.

An eloquent example of this interdisciplinary intersection between scientific and humanistic approaches to the 'self' lies in the theme of self-knowledge through the observation of one's own face, found in both neuroscientific—as in behavioural and imaging studies purporting to seek the neural correlates of the 'self' by measuring brain response to exposure to one's own face<sup>53</sup> of 64–66 of 68 of 69—and humanistic literature—such as in Wilde's construction of Dorian Gray, a character of Faustian inspiration, ii who, grappling with the transient nature of his own beauty, sells his soul in exchange for having his portrait age

in his stead. 82 In Wilde's novel, the portrait serves as an exteriorised 'self', the transposition of an immaterial entity (Dorian's inner 'self') into a physical, observable object (Dorian's picture). Interestingly, because each sin is displayed as a sign of ageing, Dorian's portrait enables him to gaze upon his own physical identity and to be visually confronted with the consequences of his own acts. In doing so, Wilde's narrative extends the literary construction of self-knowledge through visual appraisal beyond an exteriorised physical doubling of the 'self' to incorporate the objectification of the behaviour emanating from it. The visual rendition of the immorality of Dorian's behaviour in the form of exterior cues he can gaze upon raises issues of self-observation and self-monitoring as a central component of Wilde's social psychological construct of selfhood. Likewise, the transformation of actions—of which Dorian is author—into visible marks available for his inspection entails a shift from first- to third-person perspective, as Wilde's protagonist becomes at the same time subject and object of his own behaviour, actor and spectator on the stage of his immortal life.

The dual-subjective and objective-nature of the 'self' provides a further example of humanistic-scientific crossfertilisation and convergence, as it appears both (implicitly) in literary representations—such as in Wilde's narrative—and (more explicitly) in psychological accounts of selfhood—as in William James' earliest formulation of the concept of 'self' in modern psychology: 'Whatever I may be thinking of, I am always at the same time more or less aware of myself, of my personal existence. At the same time it is I who am aware; so that the total self of me, being as it were duplex, partly known and partly knower, partly object and partly subject, must have two aspects discriminated in it, of which for shortness we may call one the Me and the other the I.' (Psychology: briefer course, p. 187). 83 Interestingly, almost 300 years before Wilde and James published their insights on the 'self' as presenting us with the challenge of managing two levels of apprehension of realitythe I, subjective knower, and the Me, object that is known roman artist Artemisia Gentileschi (1593-1653) had already provided an eloquent pictorial take on the same theme: in an unconventional self-portrait, most likely achieved through a two-mirror arrangement, she represents herself as gazing not at the onlooker—as tradition would have had it—but at the canvas on which she is working, and at which the onlooker too is gazing (figure 2), epitomising the dual nature of her identity, being at the same time the subject and the object, her own artefact in a continuous state of self-creation.

Further testimony to the interdisciplinary potential of the study of selfhood emerges from the fact that, despite the difficulties in identifying a specific neural signature associated with viewing one's own face, there exists wide acceptance within the neuroscientific community of the semi-intuitive notion that one's own face is 'special', an idea closely reminiscent of the mesmerising effect attributed to the reflected face in Narcissus' myth. Evidence in favour of this singular property comes from the observation that, compared with other faces, one's own face is identified faster, 84 is primed by other self-related cues, 53 and interferes more with (ie, slows down performance in) ongoing tasks. 85 All these phenomena have been taken to be multiple manifestations of the same underlying cause: the unique ability of one's own face to capture the viewer's attention. 85–87 Evidence for special saliency of self-referential stimuli (ie, stimuli referring to the observer) has been found beyond the visual modality as well. In the auditory domain, it has long been shown that one's own name is an exceptionally attentiongrabbing cue, able to induce selective perceptual attunement to

iiBased on the German legendary character of Faust, a highly successful but unsatisfied scholar who offers his soul to the devil in exchange for unlimited knowledge and worldly pleasures (Münckler M. Melancholy and Despair: the 'Historia von D. Johann Fausten'. In: Sieber A, Wittstock A, eds. *Melancholie—zwischen Attitüde und Diskurs*. Göttingen: V & R Unipress, 2009).



**Figure 2** Artemisia Gentileschi (1593—1653). *Autoritratto in veste di Pittura (Self-Portrait as the Allegory of Painting)* (1639). Oil on canvas. London, Royal Collection, Kensington Palace. In this unconventional self-portrait, probably achieved through a two-mirror arrangement, the painter gazes at the canvas on which she is working, and in which she is represented, thus epitomising the dual nature of her 'self', both subject and object, her own artefact in a continuous state of self-creation.

previously ignored auditory channels,88 to be detected under noise-masking conditions<sup>89</sup> and to easily awaken from sleep.<sup>90</sup> Results pointing to the primacy of one's own name have been shown to extend to other domains, 91 92 as in the 'Name Letter Effect',  $^{93}$   $^{94}$  an unconscious preference for letters occurring in one's own name as compared with letters occurring in other names. Taken together, the experimental evidence available so far seems to indicate that mere belongingness to the 'self' is sufficient to enhance the attractiveness of the object of perception. I argue that by having long been present in ancient narratives—such as in Narcissus' legend—and corroborated two millennia later by neuroscientific results, the notion of the unique attractiveness one's own face represents a junction where neuroscientific and humanistic enquiry on the 'self' converge. Taking one further leap across disciplines, I propose that although, traditionally, Narcissus' tragedy has been taken to lie in his vanity and self-love, the discovery of autoprosopagnosia, a neurological condition associated with focal brain damage, causing selective failure to recognise one's own (but not someone else's) face, 95 prompts an alternative reading whereby the divine punishment consists in rendering him unable to recognise his own reflection. Interpreted in terms of loss of self-recognition ability, Narcissus' legend becomes a powerful allegory for the notion of self-non-self distinction as a key requisite for survival, thus capturing in an ancient, compelling and sophisticated narrative what is today considered a fundamental principle across virtually all biological levels, from immunology to cognition.  $^{96}$ 

# THE (INTER)FACE BETWEEN THE 'SELF' AND THE 'OTHER'

By placing the face at the core of the cognitive self—non-self distinction, Narcissus' myth resonates with influential philosophical ideas postulating the face as the locus of relationality and separation, simultaneously acting as interface and unbridgeable boundary between the 'self' and the 'other', through which the 'self' is partly defined.

The relevance of the 'other' to the debate about the 'self' recurs in philosophical thought since Greco-Roman antiquity, and is often formulated in solipsistic terms, whereby the individual is conceived as inseparable from his or her surroundings. In Plato's Republic (ca. 380 BC), 97 as well as in the Hippocratic corpus (ca. 400 BC), for instance, the soul was inextricably united with the physical world in which it was embedded. Similarly, the medieval individual existed primarily within society, be it the extended family, the parish or the feud. The concept of the 'self' as inseparably enlaced in a living world was also grounded in the animistic tradition, still widespread beyond the 16th century. 98 Likewise, in most ancient and modern eastern cultures the 'self' is inscribed into an interdependentholistic-centripetal psychosocial orientation, unlike the modern Western 'self', which mostly reflects an independent-individualistic-centrifugal attitude. Importantly, be it construed holistically or individualistically, the concept of 'self' is invariably embedded within—and partially defined by—the concept of 'other', which it intersects at a boundary whose position is constantly renegotiated based on the evolution in the notion of 'self' and 'personhood' over history and in different cultures.

The notion of face as the physical and conceptual boundary where this negotiation and mutual definition between the 'self' and the 'other' occurs is relevant to social theories of selfhood predicated on the idea—rooted in Hegelian dialectic—that mind and self emerge from the social process of communication. In Mead's pragmatistic account, <sup>100</sup> for instance, the 'self' is conceived as '[...] an eddy in the social current [...]' (Mind Self and Society, p. 182), becoming an object to itself in virtue of its social relations to others (ibid., p. 225), in which the face plays a prominent role as locus of expression and communication. The pivotal role of the face in navigating the social world and defining the 'self' constitutes the central tenet of Lévinas' theorisations of the intersubjective experience through face-toface encounter as the most affectively disrupting event for human consciousness. 101 For Lévinas, the 'I' discovers itself when it is singled out by the (simultaneously interrogative and imperative) gaze of the 'other', which, in its nudity and defenselessness, reminds the 'self' of its own 'irreplaceable responsibility towards the "other" ' (The Lévinas Reader, p. 181). 102 In being both the origin and the target of the intersubjective gaze, the face embodies the original site of the sensible, enabling the 'exceptional presentation of self by self' (Totality and Infinity, p. 202). 103 Likewise, by expressing the summon 'do not kill me' (tu ne me tueras point), Lévinas' face urges a response from which dialogue—and ultimately language—arises. Hence, by reaching to the 'other' while embodying the 'self', Lévinas' face acts as the supreme locus of relationality and separation, 'breaking through the form that nevertheless delimits it' (ibid., p. 198).

Lévinas conceptualisation of the face as expressing the dialectic between selfhood and otherness provides yet another example for interdisciplinarity in the study of the 'self', as it

# Original article

finds echoes both in ancient philosophical thought and in contemporary scientific enquiry. Aristotle (350 BC) thus comments on the etymology of the Greek word prosopon (face): 'In humans the part between the head and neck is called prosopon, a name derived, it seems, from its function. For the only animal that stands upright is also the only one that looks directly ahead and sends its voice forward' (On the Parts of Animals, Book III). 104 By tying gaze to speech, Aristotle's remark seems to imply other people's presence as partly defining the human 'self'. 105 In Kant's philosophy, the 'self', as empirical and transcendental consciousness, cannot be the object of experience without placing itself in the external world of real objects. Similarly, Wittgenstein's attribution of experiences to oneself (eg, sensations) is only possible if these experiences can be ascribed to others. 78 iii Notably, Wittgenstein's 'self' is not fully a subject either, as it 'shrinks to a point that is absent of personal identity' (as pointed out in Atkinson's The Mystical in Wittgenstein's Early Writings, p. 62). 106 This leads to a denial of the dichotomy between inner 'self' and external world (Wirklichkeit), and to a solipsistic view in which only reality filtered through subjective perception (Realität) exists (ibid., p. 63). Abandoning the assumption of a substantive ego, Wittgenstein's position seems to fulfil Schopenhauer's wish that the 'principium individuationis' be 'lifted from the eyes of a man to such an extent that he no longer makes the egoistical distinction between himself and the person of others [...] recognizing in all beings his own true and innermost self' (The World as Will and Representation, p. 378). 107 Similarly, Wittgenstein's rejection of a sharp dichotomy between inner 'self' and external reality resonates with Alva Noë's argument against the widespread brain-centric assumption—rooted in Crick's 'astonishing hypothesis'—purporting that consciousness is exclusively a product of the human brain. 'We spend all our lives embodied and situated and involved with the world around us. How could we take seriously the idea that consciousness depends only on what happens inside the brain?' (Out of Our Heads, p. 172). 108 Consciousness does not happen—Noë maintains—we achieve it, through dynamic interaction between our situated embodiment and our surroundings. Thus, rather than being private, Noë's selfhood is public, and located at the intersection between brain, environment, culture and society.

Wittgenstein's, Schopenhauer's, and Noë's stances are all relevant to current scientific research on the 'self' as, through the notion that the subjective is always confounded by the social, and the social can only be grasped from a subjective point of view, 109 they question—and ultimately reject—the viability of third-person epistemological approaches explaining selfhood in terms of physical and biological workings of the human brain, understood as a (at least in principle) totally analysable and predictable system. This scepticism towards the exhaustivity of the third-person method finds further substantiation in the consideration—advanced by the so-called 'soft line reductionists'110—that the constant and exclusive focus on external evidence carries the risk of 'forgetting'—and perhaps eventually denying—the internal subjective processes that cannot be scientifically observed. I argue that the sense of 'self' exemplifies one such elusive and yet pervasive aspect of human existence, whose complexity can hardly be satisfactorily captured exclusively in objective terms.

# CONCLUSIONS: TOWARDS A NEUROPHENOMENOLOGY OF THE 'SELF'

Neuroscientific research, by focusing on objectively measurable physical events taking place in the brain, enables the acquisition of knowledge at the proximal (how?) level. Philosophical enquiry, as well as artistic and literary production in the humanistic tradition, by constituting records of individual intuitions and experiences, rescue the subjective dimension of knowledge, which it may be able to address at the ultimate (why?) level. I argue that integrating the subjective perspective to objective methods enables retaining the explanatory power of the brain as physical system, while increasing our potential to capture emergent phenomena that escape purely physical, third-person epistemological approaches.

Building on these considerations, I propose an alternative epistemological approach to the study of selfhood: I argue that the dual (subjective and objective) nature of the 'self' uniquely positions it at the intersection between scientific and humanistic research, calling for interdisciplinary efforts embracing both the objective accounts of mainstream neuroscience, the subjective perspective of phenomenological enquiry and the intuitive insight found in artistic production, allowing all to contribute their explanatory potential, but denying each *alone* the possibility to attain exhaustive knowledge of it. I suggest that this integrated *neurophenomenological* approach, where neuroscience and the humanities merge their endeavours, can guide a more holistic research, whose methods are better equipped to tackle the complexity of selfhood.

The neuroscientific exploration of the 'self', a topic traditionally falling within the realm of philosophical enquiry, is in its infancy. As Kircher and David pointed out, <sup>17</sup> the purpose of novel enquiries in a previously unexplored domain is often that of explaining, through empirical results from a given field (eg, neuroscience), a concept developed—or a phenomenon observed—in a different field (eg, philosophy, psychology). In this sense, the study of the 'self' affords a precious opportunity for cross-disciplinary discourse between neuroscience, psychology, philosophy, medicine, anthropology and even art history. For humanists and scientists alike, studying the 'self' presents therefore an enriching opportunity to complement and challenge one's own method of inquiry, without renouncing it, with alternative approaches to knowledge, proper to other disciplines.

**Acknowledgements** The author gratefully acknowledges support through the 'Art & Neuroscience' fellowship from the Italian Academy for Advanced Studies at Columbia University. The author wishes to thank the two anonymous reviewers for insightful and constructive comments on a previous version of the manuscript.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

#### **REFERENCES**

- Encyclopedia.com. "Upanishads" The Columbia Encyclopedia. 6th edn, 2008. http://www.encyclopedia.com/doc/1E1-Upanisha.html (accessed 3 Jul 2010).
- Pausanias. Description of Greece. (English translation by JG Frazer). New York: The Macmillan Company, 1898.
- Plato. Plato in Twelve Volumes. (English translation by P Shorey). Cambridge, MA: Harvard University Press, and London: William Heinemann Ltd, 1969.
- Aeschylus. Agamemnon. (English translation by WW Goodwin). Cambridge, MA: Harvard University Press, 1906.
- Sophocles. Oedipus the King. (English translation by EDA Morshead). London: Macmillan and Co. 1885.
- 6. Descartes R. Traité de l'Homme. Paris: Angot, 1664.
- Locke J. An Essay Concerning Human Understanding. 27th edn. London: Tegg and Son, 1836:226.
- Hume D. A treatise of human nature. In: Green TH, Grose TH, eds. London: Longmans, Green and Co, 1874:P226.
- Neisser U. Five kinds of self-knowledge. Philos Psychol 1988;1:35—59.

iii Although this position, occurring in the late Wittgenstein, could appear as an abandonment of solipsism, it has been argued that in reality it is still a (large-scale) solipsism, 'in which the "self" includes all the others who make up the "we" ' (Sluga 1996, p. 400).

- Legerstee M. Mental and bodily awareness in infancy: consciousness of selfexistence. J Conscious Stud 1998;**5**:627—44.
- Damasio A. The Feeling of What Happens: Body and Emotion in the Making of Consciousness. New York: Houghton Mifflin Harcourt, 1999.
- Stuss D. Disturbance of self-awareness after frontal system damage. In: Prigatano GP, Schachter DL, eds. Awareness of Deficit After Brain Injury. New York: Oxford University Press, 1991.
- Newen A, Vogeley K. Self-representation: searching for a neural signature of selfconsciousness. Conscious Cognit 2003;12:529-43.
- Zelazo PD. The development of conscious control in childhood. Trends Cogn Sci 2004:8:12-17
- Duval S, Wicklund RA. A Theory of Objective Self Awareness. New York: Academic 15. Press, 1972.
- Morin A. Let's face it. A review of Keenan, Gallup, & Falk's book "The Face in the Mirror". Evol Psychol 2003:1:161-71.
- Kircher T, David AS. Self consciousness: an integrative approach from philosophy, psychopathology and the neurosciences. In: Kircher T, David AS, eds. The Self in Neuroscience and Psychiatry. Cambridge: Cambridge University Press, 2003.
- Bruce V, Young A. Understanding face recognition. Br J Psychol 1986;77:305-27
- Liu J, Harris A, Kanwisher N. Stages of processing in face perception: an MEG study. Nat Neurosci 2002;**5**:910—16.
- Herrmann MJ, Ehlis AC, Ellgring H, et al. Early stages (P100) of face perception in 20 humans as measured with event-related potentials (ERPs). J Neural Transm 2005;**112**:1073-81
- Mouchetant-Rostaing Y. Giard MH. Delpuech C. et al. Early signs of visual categorization for biological and non-biological stimuli in humans. Neuroreport 2000:11:2521-5
- 22. Hsiao FJ, Lin YY, Hsieh JC, et al. Oscillatory characteristics of face-evoked neuromagnetic responses. Int J Psychophysiol 2006;61:113-20.
- Hoffman EA, Haxby JV. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat Neurosci 2000;3:80-4.
- Zeki SM. Functional specialisation in the visual cortex of the rhesus monkey. 24. Nature 1978;**274**:423—8.
- Livingstone M, Hubel D. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 1988;240:740-9.
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 1997;17:4302-11.
- Hasselmo MA, Rolls ET, Baylis GC. The role of expression and identity in the faceselective responses of neurons in the temporal visual cortex of the monkey. Behav Rr Res 1989:32:203-18
- Perrett DI, Oram MW, Harries MH, et al. Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Exp Brain Res 1991;86:159-73.
- Pourtois G, Schwartz S, Seghier ML, et al. Portraits or people? Distinct representations of face identity in the human visual cortex. J Cogn Neurosci 2005·**17**·1043—57
- Sui J, Han S. Self-construal priming modulates neural substrates of self-awareness. 30 Psychol Sci 2007;18:861-6.
- Kurucz J, Feldmar G. Prosopo-affective agnosia as a symptom of cerebral organic disease. J Am Geriatr Soc 1979;27:225-30.
- Bruyer R, Laterre C, Seron X, et al. A case of prosopagnosia with some preserved covert remembrance of familiar faces. Brain Cogn 1983;2:257-84.
- 33 Eacott MJ, Heywood CA, Gross CG, et al. Visual discrimination impairments following lesions of the superior temporal sulcus are not specific for facial stimuli. Neuropsychologia 1993;31:609-19.
- Adolphs R, Damasio H, Tranel D, et al. Cortical systems for the recognition of 34 emotion in facial expressions. J Neurosci 1996:16:7678-87.
- 35. Darwin C. The Expression of the Emotions in Man and Animals. New York: Appleton & Company, 1886.
- **Gallup GG.** Chimpanzees: self-recognition. *Science* 1970;**167**:86—7.
- Amsterdam B. Mirror self-image reactions before age two. Developmental Psychobiol 1972;5:297-305.
- 38 Lewis M, Sullivan MW, Stanger C, et al. Self development and self-conscious emotions. Child Dev 1989;60:146-56.
- Gallup GG. Self-recognition in primates: a comparative approach to the bidirectional properties of consciousness. Am Psychol 1977;32:329-38.
- 40 Suarez SD, Gallup GG Jr. Self-recognition in chimpanzees and orangutans, but not gorillas. J Hum Evol 1981;10:175-88.
- Epstein R, Lanza RP, Skinner BF. "Self-awareness" in the pigeon. Science 1981;**212**:695-6.
- Reiss D, Marino L. Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. Proc Natl Acad Sci U S A 2001;98:5937-42.
- 43. Plotnik JM, de Waal FB, Reiss D. Self-recognition in an Asian elephant. Proc Natl Acad Sci U S A 2006;103:17053-7.
- Singer W. Consciousness and the binding problem. Ann N Y Acad Sci 2001:929:123-46
- Keenan JP, Gallup GG, Falk D. The Face in the Mirror: The Search for the Origins of 45 Consciousness. New York: HarperCollins/Ecco, 2003.
- Howe ML, Courage ML. On resolving the enigma of infantile amnesia. Psychol Bull 1993;**113**:305-26.
- DeGrazia D. Taking Animals Seriously. Cambridge: Cambridge University Press,

- Schilhab T. What mirror self-recognition in nonhumans can tell us about aspects of self. Biol Philos 2004;19:111-26.
- 49 Kennedy JS. The New Anthropomorphism. Cambridge: Cambridge University
- 50. Mitchell RW. Mental models of mirror self-recognition: two theories. New Ideas Psychol 1993:11:295-325
- 51. Heyes CM. Theory of mind in nonhuman primates. Behav Brain Sci 1998;**21**:101-48.
- Morin A. Right hemispheric self-awareness: a critical assessment. Conscious Cogn 2002;**112**:396-401.
- 53. Platek SM. Thomson JW. Gallup GG Jr. Cross-modal self-recognition: the role of visual, auditory, and olfactory primes. Conscious Cogn 2004;13:197-210.
- 54. Brédart S. Cross-modal facilitation is not specific to self-face recognition. Conscious Cogn 2004;13:610-12.
- 55. Calder AJ, Young AW. Self priming: a short-term benefit of repetition. Q J Exp Psychol 1996:49A:845-61.
- 56. Ovid. Metamorphoses. (English translation by S Lombardo). Indianapolis: Hackett Publishing Company, 2010.
- Spenser E. Britomart (from Books III, IV, and V of the Faery Queene, edited by ME Litchfield). Boston and London: The Athenaeum Press, 1903.
- 58. Milton J. Paradise Lost, London: Wilson, 1821.
- Walker JM. Medusa's Mirrors: Spenser, Shakespeare, Milton, and the 59. Metamorphosis of the Female Self. London: Associated University Presses, 1998.
- 60. Crick F. The Astonishing Hypothesis. New York: Touchstone, 1994
- **Keenan JP.** McCutcheon B. Freund S. et al. Left hand advantage in a self-face recognition task. Neuropsychologia 1999;37:1421-5.
- 62. Keenan JP, Freund S, Hamilton RH, et al. Hand response differences in a self-face identification task. Neuropsychologia 2000;38:1047-53.
- Keenan JP, Nelson A, O'Connor M, et al. Self-recognition and the right hemisphere. Nature 2001;409:305.
- Sugiura M, Kawashima R, Nakamura K, et al. Passive and active recognition of 64 one's own face. Neurolmage 2000;11:36-48.
- 65 Kircher TT, Senior C, Phillips ML, et al. Recognizing one's own face. Cognition
- 66. Kelley WM, Macrae CN, Wyland CL, et al. Finding the self? An event-related fMRI study. J Coan Neurosci 2002:14:785-94.
- Platek SM, Loughead JW, Gur RC, et al. Neural substrates for functionally 67. discriminating self-face from personally familiar faces. Hum Brain Map 2006:27:91-8
- Uddin LQ, Kaplan JT, Molnar-Szakacs I, et al. Self-face recognition activates a frontoparietal "mirror" network in the right hemisphere: an event-related fMRI study. Neurolmage 2005;25:926-35.
- 69 Devue C, Collette F, Balteau E, et al. Here I am: the cortical correlates of visual selfrecognition. Brain Res 2007;1143:169-82.
- Preilowski B. Self-recognition as a test of consciousness in left and right hemisphere of "split-brain" patients. Act Nerv Super (Praha) 1977;19:343-4.
- Breen N, Caine D, Coltheart M. Mirrored-self misidentification: two cases of focal onset dementia. Neurocase 2001;7:239-54.
- Turk DJ, Heatherton TF, Kelley WM, et al. Mike or me? Self-recognition in a split brain-patient. Nat Neurosci 2002;5:841-2.
- 73 Uddin LQ, Rayman J, Zaidel E. Split-brain reveals separate but equal selfrecognition in the two cerebral hemispheres. Conscious Cogn 2005;14:633-40.
- 74. Kircher T, Seiferth NY, Plewnia C, et al. Self-face recognition in schizophrenia. Schizophr Res 2007;94:264-72.
- Descartes R. Meditations on first philosophy, with selections from the objections and replies Translated and edited by Cottingham J. Cambridge: Cambridge University Press, 1996.
- Frascolla P. Understanding Wittgenstein's Tractatus. New York: Routledge, 2007. 76.
- Wittgenstein L. Notebooks 1914—1916. 2nd edn. (German text with English translation by GEM Anscombe, edited by GH von Wright, and GEM Anscombe).
- Chicago: The University of Chicago Press, 1979.

  Sluga H. "Whose house is that?" Wittgenstein on the self. In: Sluga HD, Stern DG, 78. eds. The Cambridge Companion to Wittgenstein. Cambridge: Cambridge University Press, 1996.
- Berrios GE, Markova IS. The self and psychiatry: a conceptual history. In: Kircher T, David A, eds. The Self in Neuroscience and Psychiatry. Cambridge: Cambridge University Press, 2003.
- Nagel T. What is it like to be a bat? Phil Rev 1974;4:435-50. 80.
- Zahavi D. Subjectivity and Selfhood. Cambridge, MA: The MIT Press, 2006.
- Wilde O. The Picture of Dorian Gray. Project Gutenberg, 1890. http://www. gutenberg.org/files/174/174-h/174-h.htm (accessed 25 Jan 2011)
- James W. Psychology: Briefer Course. New York: Henry Holt and Company, 1892. 83.
- Tong F, Nakayama K. Robust representations for faces: Evidence from visual search. J Exp Psychol: Hum Percept Perform 1999;25:1016-35.
- Brédart S, Delchambre M, Laureys S. One's own face is hard to ignore. Q J Exp Psychol 2006;59:46-52.
- 86. Lavie N, Ro T, Russell C. The role of perceptual load in processing distractor faces. Psychol Sci 2003;14:510-15.
- Gray HM, Ambady N, Lowenthal WT, et al. P300 as an index of attention to selfrelevant information. J Exp Soc Psychol 2004;40:216-24.
- Moray N. Attention in dichotic listening: affective cues and the influence of instruction. Q J Exp Psychol 1959;11:56-60.

# Original article

- Howarth CI, Ellis K. The relative intelligibility threshold for one's own name compared with other names. Q J Exp Psychol 1961;13:236—9.
- Oswald I, Taylor A, Treisman M. Discrimination responses to simulation during human sleep. *Brain* 1960;83:440—53.
- Snyder CR, Fromkin HL. Uniqueness: The Human Pursuit of Difference. New York: Plenum Press, 1980.
- 92. **Dion KL.** Names, identity, and self. *Names* 1983;**31**:245—57.
- Nuttin JM Jr. Narcissism beyond Gestalt and awareness: the name letter effect. Eur J Soc Psycho 1985;15:353—61.
- Hoorens V, Todorova E. The name letter effect: attachment to self or primacy of own name writing? Eur J Soc Psychol 1988;18:365—8.
- Schachter M. Paroxystic auto-prosopagnosia and epilepsy. Physio-pathological correlation or chance association? Schweiz Arch Neurol Neurochir Psychiatr 1976:119:167—76.
- Simpson SJ, Hines PJ. Self-discrimination, a life and death issue. Science 2002;296:297.
- Plato. The Republic. (English translation by B Jowett). Oxford: Oxford University Press, 1892.
- Harvey G. Animism: Respecting the Living World. New York: Columbia University Press, 2006.
- Chang SC. The self: a nodal issue in culture and psyche—an Eastern perspective. *Am J Psychother* 1982;36:67—81.

- Mead GH, Morris CW. Mind self and society from the standpoint of a social behaviorist. Chicago: University of Chicago Press, 1934.
- Lévinas E. Emmanuel Lévinas: basic philosophical writings. In: Peperzak AT, Critchley S, Bernasconi R, eds. Bloomington: Indiana University Press, 1996.
- 02. Lévinas E. The Lévinas reader. In: Hand S, ed. Oxford: Blackwell Publishers, 1989.
- Lévinas E. Totality and Infinity: An Essay on Exteriority. Dordrecht: Kluwer Academic Publishers, 1991.
- Aristotle. On the Parts of Animals. (English translation by W Ogle). Oxford Editions, Oxford. 1912.
  - Reiss TJ. Mirages of the Selfe: Patterns of Personhood in Ancient and Early Modern Europe. Stanford, CA: Stanford University Press, 2003.
- Atkinson JR. The Mystical in Wittgenstein's Early Writings. New York: Routledge/ Taylor & Francis, 2009.
- Schopenhauer A. The World as Will and Representation. Vol I (Translated by EFJ Paynes, first published in 1958 by The Falcon Wing's Press). Mineola, NY: Dover Publications, 1966.
- Noë A. Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness. New York: Hill & Wang / Farrar Straus & Giroux, 2009.
- Saukko P. Studying the self: from the subjective and the social to personal and political dialogues. *Qual Res* 2002;2:244–63.
- Chalmers DJ. The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press, 1996.

# Poem

# Eat more fatty fish

and exercise more. this is what I advise my patients when their good cholesterol is low triglycerides high.

I use a letter a thin piece of paper light on words which makes for quite a lean envelope mostly bones really.

most patients swim past. and yet a few always take the bait ending up in front of crumbled ice and prices per pound.

dressed in workout clothes they ask the man in the plastic smock for a pound of fatty fish—salmon a piece freshly caught.

# **Adam Possner**

Correspondence to Assistant Professor Adam Possner, General Internal Medicine, George Washington University, Medical Faculty Associates, 2150 Pennsylvania Ave NW, Suite 2-105 South, Washington, DC 20037, USA; apossner@mfa.gwu.edu

Competing interests None.

Provenance and peer review Not commissioned; not externally peer reviewed.

Published Online First 16 February 2011

Med Humanit 2011;37:122. doi:10.1136/jmh.2011.007047