

# The regress-stopping role of affective valence

Bryony Pierce<sup>1</sup>

1

**Abstract.** This paper argues that affective phenomenology plays an essential role in reasoning, grounding reasons for action and rendering evaluative concepts subjectively meaningful. Experience of affective responses to commodities, situations and anticipated potential outcomes allows goal-setting to track changing physiological changes indicative of our current needs so that behaviour can be flexible and adaptive. Values attributed as a result of affective responses are stored and relied upon in subsequent reasoning until updated by new information conveyed, again, through affective experience. The qualitative nature of affective experience grounds reasons for action, performing a regress-stopping role, that is, stopping a possible infinite regress of whys when justifying actions or decisions to oneself. This is a distinct type of grounding from the grounding in biological utility Dickinson & Balleine's Hedonic Interface Theory claims affective experience provides; it concerns the subjective sense one has of being justified in pursuing a course of action, rather than offering an evolutionary explanation of the origins of the motivation to act. It is also distinct from the kind of grounding discussed by Smithies or Chudnoff in relying on affective content rather than perceptual content or cognitive phenomenology.

## 1 GROUNDING REASONS FOR ACTION

We think of our reasons as grounded, from a first-person perspective and when contemplating other people's reasons for action. That is to say that certain facts concerning reasons are taken to obtain in virtue of something more fundamental, in a relation of non-causally-dependent justification. I will argue that conscious experience of affective valence plays an important regress-stopping role in (human) action production.

According to Dickinson & Balleine's Hedonic Interface Theory (HIT), the function of consciousness is to act as an interface between cognition and emotion, enabling goal-directed action [1,8]. Affective responses are key in their account. They continually update values, enabling goals to be set that are responsive to changing physiological needs in a changing environment, so that behaviour can be flexible and adaptive, unlike stimulus-response behaviour. HIT's claim is, however, qualified to include the possibility that some functional equivalent of consciousness could have enabled goal-directed action. But the need for reasons for action to be grounded, from a subjective point of view, so that they make sense to us, allows a stronger claim to be made: that consciousness itself is necessary for goal-directed action that can be judged (subjectively) to be performed for reasons, by those able to explicitly evaluate goals. This claim is about human goal-

directed action, though I do not rule out the possibility of a similar role for consciousness in some other animals.

When speaking of a functional equivalent of HIT's interface, one important function is allowing information about felt emotional states to be accessible to reasoning. It is not enough for action production to be considered merely in behavioural terms, and so I reject the idea that some (non-conscious) functional equivalent of consciousness, such as a non-conscious transducer, could ground value during the reasoning process, as Dickinson suggests could have been the case [8, p. 78]. This is because, in humans, at least, consciousness performs a regress-stopping function.

I offer two arguments in support of my claim that this function cannot be fulfilled by a non-conscious transducer:

1. The way the interface would process information during reasoning would render the functional equivalent of consciousness indistinguishable from consciousness, and so it would qualify as consciousness.
2. A non-conscious transducer cannot confer meaningfulness to information and thus lend perceptible motivational force to (abstract) reasons.

Argument (1) relies on the premise that, were the functional criteria to be met, those for being a subject of conscious experience would also be satisfied. The functional criteria, in my account, include the requirement that subjects be capable of recognising reasons as meaningful and sufficient to justify action. Subjects must be able to assess different courses of action on the basis of their value, so that they are able to form preferences capable of motivating action, and, in my account, possession of the concept of value requires the capacity for processing information about affective experience. A functional equivalent of consciousness would need to be able to allow judgements about relative value to be made on the basis of affective experience, and emotions would need to be represented in abstract form in order to be able to feature in reasoning. Subjects would both possess a concept of relative value and experience affective responses with motivational force, which they could report, so it would seem to them that they were conscious: their mental states would be indistinguishable from consciousness as we understand it.

From the first-person perspective, a functional equivalent of consciousness, as specified above, would meet the criteria for being conscious in virtue of its indistinguishability from consciousness *simpliciter*. In this case, I argue that it would *be* consciousness. Dennett questions whether there seeming to be phenomenology is sufficient for there really to be phenomenology [7, p. 366], but I agree with Strawson [20, p. 113] that, "for there to seem to be phenomenology is for there to be phenomenology" and that in the case of experience, we "can't open up the is/seems gap". For there to seem to be experience of affective responses with the relevant function is for there to be something it is like from a first-person perspective that, since it is indistinguishable from conscious experience of those

<sup>1</sup> Dept. of Philosophy, Univ. of Bristol, BS, UK. Email: bryony.pierce@bristol.ac.uk.

responses, must qualify as conscious experience. Anything less than seeming to experience affective responses consciously would fail to be a functional equivalent of consciousness – it would be mere (ungrounded) information processing, as in a computer program that reported responses without access to the underlying meaning. From a third-person perspective, the two cases (subjects with consciousness or a functional equivalent of consciousness) would be indistinguishable, if the subject reported having conscious affective experiences. There would be scope for scepticism about whether subjects really were conscious, but no more so in the case of reports from those with a functional equivalent of consciousness than from those who were conscious, all other things being equal, as the phenomenal character of affective experience has a function, so features in either case. I say more to defend this claim, and argument (2), below.

## 2 DIFFERENT WAYS OF GROUNDING REASONS

I will now distinguish the regress-stopping sense of grounding reasons from other ways of grounding reasons. I want to start by making a distinction between grounding in the sense of being explicable in terms of a biological or other function and grounding in the sense of having meaning, as in Searle's Chinese room thought experiment [17].<sup>2</sup> The Chinese room argument distinguishes between the ability to manipulate symbols according to syntactic rules and the ability to understand their meaning in the sense of their semantic content. I make a further distinction between semantic content in the sense of reference to that which is external to the symbol system and the subjective meaningfulness of information – its affective content. The word 'home', for example, can simply refer to the building that is one's residence, or convey also the emotional relation in which one stands to a place, a feeling of ease or attachment, perhaps. Affective content plays an important role in my argument about grounding reasons for action.

When we say that reasons or underlying desires are grounded in biological states, we are explaining their function within a biological system or telling an evolutionary story about why they exist or are useful. In HIT, consciousness is necessary to communicate information about physiological states to cognition, so that it is available in means-end reasoning. "desires must... be grounded in something that the agent cares about", the desires of any animal 'must, ultimately, be grounded in its biology' [1, p. 81], and agents only know what to desire once the effects of events on physiological states are made explicit through conscious affective responses. The role of these responses is key, yet reasons are *grounded*, for Dickinson and Balleine, in underlying biological functions. This is one type of grounding, but if we are to find reasons for action compelling, another type of grounding is necessary: grounding in the affective valence of subjective experience.

<sup>2</sup> In Searle's thought experiment, a monoglot English-speaking person follows instructions in English to manipulate Chinese symbols to produce correct responses to questions in Chinese, but Searle argues that this does not show that the English-speaking person thereby understands Chinese, or, by analogy, that a computer understands the strings of symbols it manipulates, even if it can simulate understanding.

There are two main problems in Dickinson and Balleine's account that I will address specifically:

1. *The problem of meaningfulness*: the level at which grounding takes place is in physiological states, and these lack any intrinsic meaning.
2. *The problem of circularity*: "conscious experience" is presupposed, and although they address this by weakening their claim so that it is only consciousness *or some functional equivalent* that is said to be necessary for goal-directed action, this solution is unsatisfactory.

I will return to the problem of circularity after introducing and discussing the problem of meaningfulness. The fact that conscious experience might have adaptive value for a biological organism and the question of how symbols can refer to anything external to the symbol system are only indirectly related to how a person can justify reasons, from a subjective point of view. I will examine the role meaningfulness, in the sense of affective content, plays in grounding reasons for action, that is, its significance when defending the claim that consciousness has a regress-stopping function.

Harnad takes the symbol grounding problem to be the problem of how we can move from abstract symbolic representations, such as words describing objects, to the non-symbolic (the objects described) in a way that stops the infinite regress of explanations of the meaning of symbols in terms of other intrinsically meaningless symbols [9]. Although concerned with meaning, it is distinct from the problem of *meaningfulness*: the problem of how symbols can have subjective meaning in terms of incorporating motivational force. Harnad proposed a solution to the symbol grounding problem by grounding meaning in non-symbolic sensorimotor function [10, 7.2]:

Hence grounding means sensorimotor grounding: Symbols must be grounded in the capacity to discriminate and identify the objects, events and states of affairs that they stand for, from their sensory projections

This solution provides information with a functional application, and makes the move from a symbolic form to some kind of physical instantiation, in the form of the relation between sensorimotor activity and objects of perception. An important further question, in the context of the role of consciousness in goal-directed action, is how information-carrying strings of abstract symbols taken to be reasons for action acquire meaningfulness to the agent, such that the agent deems them to be compelling and feels justified in acting upon them. This meaningfulness is more than a relation to biological utility or to something external to the system of symbols, although it involves both. It additionally provides compelling motivational force in the form of goals that can be judged to have self-evident value. The meaningfulness of the goal, which depends on its perceived value, in a particular context, provides subjective reasons for action, which I contend can have normative as well as motivational force.

It is necessary to ground symbols, in the case of symbols representing reasons,<sup>3</sup> by relating them to objects that are external to the symbol (or other) system, to stop the infinite

<sup>3</sup> This account applies when seeking to ground reasons regardless of whether a symbolic or representational theory is presented – it could also apply to dynamical systems that are taken to be non-symbolic, for example, as the need to ground reasons remains. I will not go into whether these might also be considered symbolic, as nothing important to my claim rests upon the nature of the implementation.

regress created by using symbols to define symbols (or processes to explain the value of other processes). The problem is that doing so fails to address the question of what halts the regress of whys so as to justify action to the agent. At some point we need to be able to stop asking why we ought (prudentially or morally) to do something, because it becomes self-evident that a reason justifies something, non-inferentially, without further explanation. (This reason need not be entertained consciously at the time of deliberation; it may be recognised tacitly.)

Harnad's solution of grounding symbols in sensorimotor activity is one type of grounding that aims to address Searle's concern [17], but it doesn't go far enough if we want to justify reasons to ourselves, within our reasoning, because sensorimotor activity isn't an end in itself.<sup>4</sup> If we cannot proceed on the basis of value-based premises, set goals, and judge our own reasons to be justified, we cannot see them as good or compelling reasons to pursue goals. So, we need to be able to attach value to the *consequences* of sensorimotor activity. The proposal to ground these consequences in biological utility, although consistent with an evolutionary explanation of how we came to have certain dispositions, doesn't stop the regress of whys when reasoning. It is always possible to question why one's biological survival (or the replication of one's genes) should be viewed as having value, and the answer is not self-evident. The futility of life is a theme in the Arts (e.g. the Theatre of the Absurd) and there are many groups that promote other values more highly than the biological survival of oneself and one's kin (e.g. religious, terrorist or other political organisations). Societies often promote the value of 'immortality' through various achievements that live on beyond the individual (e.g. works of art; acts of self-sacrifice; philanthropic acts), which may do nothing to further that individual's interests biologically.

My aim is to identify a way of grounding reasons subjectively, to oneself, when making judgements, which I argue is made possible by the representation of physiological states in the form of affective responses. Grounding occurs, not at the physiological level, in this case, but at the level of abstract representation of affect. Feelings and emotions are intrinsically value-laden; they have hedonic tone and a positive or negative valence. I claim that reasons for action, desires and subjective meaning are grounded, in the relevant sense, in emotion: affective responses lend meaning to the abstract concepts of 'good' and 'bad', and make evaluation possible. From an evolutionary perspective, affective experience can also be said to be grounded in the biological utility of monitoring and adjusting the underlying physiological states emotions represent, if we accept something like the James-Lange theory of emotions [13,14].<sup>5</sup> Evolutionary accounts address the explanation of why things are the way they are, but fail to address *the problem of meaningfulness*.

I am not claiming that the phenomenal character of affective responses is at a more fundamental level than biological function, but that it is appropriate to ground reasons for action in this way, alongside grounding in biology. This allows us to

address the need for reasons to have subjective meaningfulness, in order to have motivational force and prompt action.

Teleological theories, in philosophy [15], aim to explain how mental representations can represent objects or states of affairs in the external world. Mental content is grounded in biological functions selected by natural selection. Teleosemantic theories are teleological theories that seek to offer a naturalistic explanation of the semantic properties of representations (*ibid.*). Meaning, like mental content, within such a theory, is explicable in terms of (and so grounded in) adaptive functions, with the understanding of meaning involving no more than having some minimal knowledge of the referent of a word. My account of grounding is distinct from those proposing a mapping from symbols to physical objects or states of affairs in the external world. In my account symbols are mapped onto abstract objects (good, bad, beautiful, frightening, etc.), which are inaccessible via sensorimotor activity, and which are not intrinsic to adaptive behaviour (adaptive behaviour in plants does not require the representation of abstract objects).

The *symbol grounding problem* in the field of artificial intelligence (AI), and in philosophy seeks to explain the related problem of how abstract symbols can refer to anything external to the symbol system. An element common to many accounts of grounding (including HIT, teleological accounts and Harnad's account) is that there should be a causally efficacious realiser – the information processed must have some means of acting upon the external world. HIT explicitly incorporates the motivation, as well as the capacity, to act on the external world. Whereas teleological and teleosemantic theories propose explanations that appeal to survival- and reproduction-related purposes; the symbol grounding problem focuses on how reference and interaction could be possible.

Another approach to grounding reasons, in the philosophy of action literature, examines whether reasons for action are grounded in desires, beliefs or values (see [6]), distinguishing between motivating and normative reasons – between what we desire and what we ought to do, rationally or morally. The approach I take is differentiated from this one by its focus on grounding in the sense of being able to justify reasons for action *to oneself*.

Turning to recent work in epistemology, Smithies says that “the problem of explaining mental representation cannot be divorced from the problem of explaining phenomenal consciousness” [19, p. 98]. Smithies distinguishes between appeals to phenomenal consciousness and to functional role, arguing that phenomenal consciousness provides immediate non-inferential justification for beliefs. In contradistinction to his position, I argue that consciousness has a (compound) functional role, and that conscious affective responses provide immediate non-inferential justification for reasons for action. He uses the example of blindsight<sup>6</sup> to illustrate how phenomenal rather than functional differences play a crucial role in epistemic justification: when the phenomenal character of perceptual experience is absent, blindsighted subjects lack epistemic justification, even though information is being processed unconsciously and they may reliably distinguish objects and orientations in forced-choice experimental situations. Smithies argues that phenomenal consciousness is the basis for epistemic justification, and that mental representation *that plays an*

<sup>4</sup> Harnad has since acknowledged that “What makes the hard problem hard is precisely the mysterious difficulty of explaining feelings functionally. So the mind-body problem is actually the ‘feeling-function’ problem” [11].

<sup>5</sup> Prinz, for example, defends a modified view, arguing that “emotions are somatic, but they are also fundamentally semantic” [16].

<sup>6</sup> Blindsight subjects can respond to visual stimuli, but lack conscious visual experience (see [22]).

*epistemic role* is grounded in phenomenal consciousness. Smithies restricts his claim to epistemic justification. A further significant difference between our accounts is that for Smithies grounding is in conscious perceptual content, not affective content. What our accounts have in common is that grounding is in conscious experience. Although I specify that it is consciousness of affective responses that plays a crucial role in grounding, my claim encompasses consciousness in general, as on my account no phenomenally conscious experience is entirely devoid of affective content. This is a consequence of consciousness's role as an interface between emotion and cognition; within that interactive process, affective valence, which is one important aspect of affective experience (affective experiences differ in ways other than simply in terms of valence), is an ever-present integral part of conscious experience.

Chudnoff also addresses epistemic justification, arguing that it is in virtue of their phenomenology that intuitions justify beliefs [4]. For Chudnoff, the phenomenology described ("presentational phenomenology") has perceptual content, but it can also have intuition-based content [3,4]. The basis for epistemic justification is then the phenomenal character of perceptual or intuition experience. The justification provided by perception is *prima facie* – "it can be defeated or undermined" [3, p. 315], just as reasons for action that seem to be justified can be overturned. Perceptual presentational phenomenology consists in there being perception of both a fact and an item in virtue of which this fact is perceived to be the case. In the case of intuition experience, which may be imaginative or cognitive, one fact-intuits *p* and is intellectually aware of some item that makes it the case that *p*.

Chudnoff says that 'presentational phenomenology can be found in cases of intuition, introspection, imagination, and recollection', but makes no mention of emotion [5, p. 51]. But his arguments do extend to emotion, and furthermore there is an affective component in conscious perception and intuition, which may also be involved in epistemic justification, and there is an evaluative component in the processes, for which affective valence is necessary. Reasons for action, as well as beliefs relating to perceptual content and logical or philosophical intuitions, can have presentational phenomenology: one can come to realise what one has reason to do, upon experiencing an affective response in virtue of which a reason for action is perceived to be justified. Chudnoff's focus is exclusively on *cognitive* phenomenology, but affective responses have the same kind of immediacy/non-inferential status as intuitions, as he describes them, and intuitions have affective as well as presentational content.

### 3 HUMANS' VERSUS ROBOTS' REASONS

There is an interesting difference between the way we view a person as capable of acting for reasons and the way we view a machine or robot when it acts, and this, I argue, is not related to whether human action is free, but to the way in which reasons for action are grounded.

Whereas a simple robot performing a fixed range of tasks could follow procedures in a logical way to achieve set objectives, with appropriate programming, there seems to be no way that the robot could justify carrying out the necessary steps, providing normative reasons, without appealing to the values or objectives of the programmer or of other non-robots. Any values

or rules it uses to guide its behaviour depend on the values and objectives of those responsible for its initial programming (even if this results in further programming and learning by the robot), and the effectiveness of their implementation. In taking anything to be desirable, necessary or the correct thing to do, the robot would merely be carrying out instructions set for reasons grounded in something beyond its own programs or hardware. These reasons would not be meaningful to the robot. It has no normative reason of its own to act, just as a washing machine drum has no normative reason of its own to spin. It has no affective responses to outcomes, and this is an important reason for viewing robots and conscious entities as different (and not just with regard to whether they are moral subjects).

A robot devoid of felt affective responses<sup>7</sup> would have no normative reasons to act, let alone to perform one movement in preference to another. Any reason given, such as to help others or to maintain its own functions, would lead ultimately, when reasons were sought for each response, to the way it had been programmed, so would be grounded in something extrinsic to the robot and, being devoid of meaningfulness, would be beyond its comprehension. Taking an apparently strong reason for action, such as preventing the destruction of the planet, or any case in which failing to perform an action would cause harm or suffering, the robot might be said to have reason to act *to prevent harm or suffering*, whether self- or other-involving. Yet it seems to have no reason to prefer that nothing be harmed or that no one suffer, and no way of judging harm or suffering to be bad or undesirable that is not contingent upon a human or other entity's having provided it with that information. Although its representations (or systems) may enable it to act upon the external world, satisfying the need for a causally efficacious realiser, as in Harnad's grounding in sensorimotor function [10], they lack meaningfulness for the robot. When criteria for making value judgements are introduced by programmers, instructions may be followed blindly, as in Searle's Chinese room thought experiment [17], but the robot has no capacity to evaluate the criteria and no way of judging whether the information about value is correct or relevant to its interests. The robot's reasons are grounded in the reasons and values of others, which must in turn be grounded in their affective responses.

### 4 THE REGRESS-STOPPING ROLE OF AFFECTIVE VALENCE

Biological utility fails to stop the infinite regress of whys with respect to rational justification, because we can still question the value of our own survival and that of the replication of our genes. Affective states, even those that are potential rather than actual, succeed in stopping the regress of whys, when seeking to justify our own actions, by presenting us with information about value in a form that, at a subjective level, requires no further grounding – an immediate, non-inferential justification of action. Although we may sometimes forgo pleasure or endure pain willingly, to achieve some higher objective, our understanding of value is derived directly from affective responses to actual and

<sup>7</sup> Robots currently being developed to have or respond to emotions, such as Softbank, Aldebaran and Yoshimoto Robotics Laboratory's 'Pepper', which its developers claim has emotions, merely simulate [responses to] emotional behaviour with a view to satisfying consumer needs.

potential states of affairs. We cannot fail to believe that certain affective states are preferable to others, as it is part of their nature to have a positive or negative valence (varying in degree and quality). The feature of affective states that grounds reasons subjectively is not biological utility, though they may have it, but the valence we automatically and unquestioningly attach to them, which is not only compelling (motivating action), but is meaningful to us as thinking beings and capable of justifying reasons for action from a subjective perspective.

A human, even one convinced of the impossibility of freedom of action, can be said to have reasons for preferring one course of action over another. This requires further explanation, within a context aiming for consistency with hard determinism, as the human being is just another physical object and/or process, like a robot or washing-machine spin cycle. The ability to have reasons for preferences is explicable, to a conscious being (whether anything can be 'explicable' to a non-conscious entity is questionable, but explicability to ourselves is what is relevant here), only when we include the phenomenal quality of affective experience in our account. This is how we stop the infinite regress of whys. A human can have reason to seek pleasure, avoid pain, or act in a way that would be conducive to her own or others' wellbeing, which can be accounted for in terms of the qualities of the resultant conscious affective experiences.

Beliefs, such as the belief that pain is bad/undesirable or that eating when hungry is good/desirable, are not something we need to learn from others; it is enough to experience pain or to eat when hungry to acquire the motivation to make it a priority to avoid pain and to eat when hungry. There is something it is like, affectively, with a positive or negative valence, to experience these things. When setting objectives, we are guided by the anticipated positive or negative valence of the probable outcomes, and Dickinson and Balleine's experimental work supports the hypothesis that we update values ascribed to commodities and states of affairs only through phenomenal experience of them.<sup>8</sup> Conscious awareness of the immediate effects of actions enables perceived value to be adjusted, and the apparent implications of actions, including their potential effects on future conscious experience, guide reasoning. These effects can be judged good or bad, and desirable or undesirable, relative to the individual's values and needs (which need not track biological utility), and these judgements are grounded in regress-stopping affective valence. When they are judged good or bad in a more abstract, purportedly objective, moral or prudential viewpoint, these judgements are, I claim, also grounded in the individual's conscious experience.

Consciousness is equated here with the functional properties of an interface capable of processing information in abstract form in such a way that reasons for action can be perceived to justify action. It might be objected that this introduces a non-physical feature, a kind of property that according to some is not reducible to physical processes (see Chalmers's property dualism [2], or Jackson's knowledge argument [12]). I include (meaningful) representations of information in abstract form and subjectivity in my account, neither of which are readily reducible to the physical. But it is not inconsistent with naturalism to accept that abstract information is represented consciously or

that there is a subjective, qualitative aspect to consciousness.<sup>9</sup> Regardless of the precise ontological nature of consciousness, it is widely accepted that humans have conscious experience.

Whereas the grounding of abstract symbols must be in something *extrinsic* to the symbol system, the capacity for making judgements about value must be grounded in something *intrinsic* to the decision-maker. Reasons relying on these judgements must be such that the decision-maker requires no further justification. Reasons are grounded not only in concrete biological states, but in actual and hypothetical affective responses to actual and potential states of affairs. Awareness of the qualitative nature of these responses is necessary for information about valence to be subjectively meaningful. A person cannot judge that she will be content or distressed without any capacity for experiencing and representing what it is like to be in a state of contentment or distress. Whether or not it might be possible (logically or otherwise) to replicate this in a robot or other inanimate system is difficult to say, but it is uncontroversial to say that humans have this capacity.

Torrance, summing up Shanahan's view, said, "to talk about a being as phenomenally conscious is to talk about things mattering to that being in a way that they can't to a creature that lacks phenomenal consciousness" [21] (see also [18]). The capacity to perceive things as mattering (consciously experienced affective valence) provides the motivation to act, but it also enables us to judge some behaviour to be more rational than other alternative courses of action, because of the relative value of different potential outcomes. The pursuit of that which matters, in the sense of being associated with positive affective valence, is automatically endorsed as a motive, and that which matters is recognised, in abstract form, when reasoning, as a valid goal, grounding reasons for action. The grounding of reasons and values in affective valence thus halts the otherwise potentially infinite regress of whys, when deliberating over alternative courses of action that appear to be open to us.

## REFERENCES

- [1] B. Balleine & A. Dickinson. 'Consciousness—the interface between affect and cognition', in J. Cornwell (ed.), *Consciousness and Human Identity*, New York: Oxford University Press, 57-85 (1998).
- [2] D. Chalmers. *The Conscious Mind: In search of a fundamental theory*, Oxford: Oxford University Press (1996).
- [3] E. Chudnoff. 'The nature of intuitive justification', *Philosophical Studies*, 126 (3), 347-73 (2011).
- [4] — 'What should a theory of knowledge do?', *Dialectica*, 65(4): 561-79 (2011).
- [5] — 'Presentational phenomenology', in Miguens & Preyer (Eds.), *Consciousness and Subjectivity*, Ontos Verlag, 51-72 (2012).
- [6] J. Dancy. *Practical Reality*, New York: Oxford University Press (2000).
- [7] D. Dennett. *Consciousness Explained*, London: Penguin (1991).
- [8] A. Dickinson. 'Hedonics: The cognitive-motivational interface', in M.L. Kringelbach & K.C. Berridge (eds.), *Pleasures of the Brain*, New York: Oxford University Press, 74-84 (2010).
- [9] S. Harnad. 'The symbol grounding problem', *Physica D: Nonlinear Phenomena*: 42(1-3): 335-346 (1990).
- [10] — 'Grounding symbols in the analog world with neural nets', *Think (special issue on connectionism versus symbolism)*, 2(1): 12-78 (1993).
- [11] — 'No easy way out', *The Sciences*, 41(2): 36-42 (2001).

<sup>8</sup> Experience of things that appear to be similar or to overlap in some way may also influence value judgements, if the anticipated outcome is expected to have some of the same properties, i.e. some generalisation occurs.

<sup>9</sup> Explaining how this can arise out of physical matter in animals, not machines, is beyond the scope of this paper.

- [12] F. Jackson. 'What Mary didn't know', *The Journal of Philosophy* **83**(5): 291-295 (1986).
- [13] W. James. 'What is emotion?', *Mind*, **9**: 188-205 (1884).
- [14] C. Lange. *Über Gemütsbewegungen*, Leipzig: Verlag von Theodor Thomas (1887).
- [15] R.G. Millikan. *Language, Thought, and Other Biological Categories*, Cambridge, Massachusetts, USA: MIT Press (1984).
- [16] J. Prinz. 'Embodied emotions'. In Solomon, R. (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*, New York: Oxford University Press, pp. 44-60 (2004).
- [17] John R. Searle. 'Minds, brains, and programs', *Behavioral and Brain Sciences* **3**(3), 417-457 (1980).
- [18] M. Shanahan. *Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Minds*, Oxford: Oxford University Press (2010).
- [19] D. Smithies. 'The phenomenal basis of epistemic justification'. In Sprevak, M. and Kallestrup, J. (eds.), *New Waves in Philosophy of Mind*, Springer (2014).
- [20] G. Strawson. 'Real Naturalism', *Metodo. International Studies in Phenomenology and Philosophy*, **1**(2): 101-125 (2013).
- [21] S. Torrance. 'The "drop-out question" and the emergence of artificial consciousness', keynote presentation, 3<sup>rd</sup> AISB Workshop on the Emergence of Consciousness, London, UK, 9 May 2013, retrieved 31 July 2013 (2013).
- [22] L. Weiskrantz. *Blindsight: A case Study and Implications*, Oxford: Oxford University Press (1986).