Mirror Self Recognition and Self-Identification

Alexandria Boyle
University of Cambridge

That great apes are the only primates to recognise their reflections is often taken to show that they are self-aware – however, there has been much recent debate about whether the self-awareness in question is psychological or bodily self-awareness. This paper argues that whilst self-recognition does not require psychological self-awareness, to claim that it requires only bodily self-awareness would leave something out. That is that self-recognition requires 'objective self-awareness' – the capacity for first person thoughts like 'that's me', which involve self-identification and so are vulnerable to error through misidentification. This objective self-awareness is distinct from bodily or psychological self-awareness, requires cognitive sophistication and provides the beginnings of a more conceptual self-representation which might play a role in planning, mental time travel and theory of mind.

1. Introduction

Only some animals are able to recognise their reflections in a mirror. Within the primate order, this capacity is confined to the great apes. This is commonly taken to reflect an important cognitive difference between great apes and other primates – only the great apes are 'self-aware'. This paper addresses whether this is true, and if so, what the relevant sense of 'self-awareness' is. This has been a divisive question since Gordon Gallup developed the 'mark test' over forty years ago. In the test, an animal is marked in a location visible only in the mirror. Touching the mark when looking in the mirror is an indication that they have recognised their reflection.

Gallup takes success in the test to reveal that an animal is psychologically self-aware – that is, they have an understanding of their psychological properties. Critics argue that all that is required is bodily self-awareness – an understanding of one's bodily properties. In this paper, I argue that whilst Gallup's critics are right to deny that mirror self-recognition requires psychological self-awareness, accounts that claim mere bodily self-awareness is sufficient leave something out. Framing the debate in terms of bodily and psychological self-awareness has obscured the central problem posed by the mark test – that success requires self-identification, i.e. entertaining a thought like 'that's me', or 'I'm the thing in the mirror'. My claim here is not simply that self-recognition requires first-person thought – that is, thought whose natural linguistic expression would involve first-person pronouns. Of course it does, but it requires more than this. There is a kind of first-person thought that can unproblematically be ascribed to most animals, which I call 'subjective'. However self-identification, a thought in which one identifies oneself with an object, is of a different kind – it is objective, in the sense that it involves thinking about oneself as an object in the world.[1]

The capacity for self-identification is not simply a kind of bodily or psychological self-awareness, since a thought like 'that's me' might be motivated by an appreciation of one's psychological or bodily characteristics, but its content need not include explicit

---

[1] Throughout this paper, I use 'mirror self-recognition' and 'self-recognition' interchangeably. However, 'self-identification' as I use it is not equivalent to 'self-recognition' – since self-identification can occur in contexts other than self-recognition.

representation of either. What is distinctive in self-identification is thinking about oneself *objectively*. Consequently, I argue that mirror self-recognition demonstrates that a creature is objectively self-aware.

In §2, I describe briefly the different ways in which primates can use mirrors, concluding that an account of mirror self-recognition must explain how it differs from other mirror-involving activities, such as using a mirror to find a hidden object. In §3, I outline Gallup's 'self-concept' account of mirror self-recognition, as well as Stephane Savanah's (2013) recent defence of a similar view, and argue that these accounts are unsuccessful. In §4, I outline the two principal deflationary accounts of self-recognition – the kinaesthetic-visual matching account, elaborated by Robert Mitchell (1993, 1997, 2002, 2013), and the secondary representation account, defended most recently by Thomas Suddendorf and David Butler (2013). Whilst there are points of disagreement between these accounts, both suggest that self-recognition requires knowledge of one's physical appearance – a form of bodily self-awareness. However, I argue that this fails to make sense of the fact that monkeys do not recognise their reflections. In §5, I argue that self-recognition requires 'objective self-awareness', which is distinct from bodily self-awareness, and argue that both the deflationary accounts discussed must explicitly appeal to this fact if they are to fully explain self-recognition. I further argue that objective self-awareness involves cognitive sophistication, because it involves a relatively flexible and general representation of oneself. This in turn may form the basis of a self-concept which can be used in planning, mental time travel and theory of mind.

2. Primates and Mirrors

The mark test begins by exposing test subjects to a mirror for a period of several days. After this period, the subjects are surreptitiously marked with a dye on one eyebrow ridge and one ear, in such a fashion that the marks will be visible to the subject only in the mirror. After this, the subjects are monitored in the absence of the mirror to determine the number of times they touch the marks. The mirror is then reintroduced, and mark touching is monitored again. A subject 'passes' the mark test if mark touches during the mirror-present phase significantly outnumber mark touches during the mirror-absent phase.

If the mark test is conceived of as a test of self-recognition, it can present false positives (Mitchell, 1993). A subject might touch its own mark without self-recognising – perhaps thinking 'that guy's marked – maybe I am too?' But the mark test is intended to formalise and support other behavioural criteria for self-recognition (Gallup, Anderson, & Platek, 2011, p. 91). For instance, the prevalence of 'self-directed behaviour', such as making faces in the mirror and using the mirror to inspect body parts visible only in the mirror can also be measured. As Mitchell (2012) argues, the relationship between these behavioural criteria is complex – it appears that they can be met independently of one another. Consequently, it is not clear precisely what criteria should be used to decide whether an animal is capable of self-recognition. What should be uncontroversial given the evidence, is that even on a relatively strict set of criteria for self-recognition, some great apes recognise their reflections.

Among primates, there is compelling evidence of self-recognition only in the great apes (Suddendorf & Butler, 2013).[2] This reveals a striking difference between great apes

---

[2] A recent study found that macaques pass the mark test and display some self-directed behaviour after extensive training (Chang et al., 2015). However, whether this constitutes evidence of self-recognition in monkeys has been contested by Anderson and Gallup (Anderson & Gallup, 2015), who suggest that the macaques' behaviours are a mere

and other primates. Despite this, monkeys have demonstrated a range of other competences with mirrors. For instance, they are able to use mirrors to locate a variety of objects (Anderson, 1986; Itakura, 1987b), from food items to images of other monkeys. This ability to locate mirrored objects has been distinguished from 'mirror guided movement' – the ability to use the mirror to visually guide the movements of the limbs when reaching for objects, where the reaching limb is visible only using the mirror (Menzel, Savage-Rumbaugh, & Lawson, 1985). Shoji Itakura (1987a) found that Japanese macaques can learn to do this, in situations ranging from selecting a box containing a piece of food to pressing a sequence of illuminated keys in the correct order, in exchange for a food reward. Despite these abilities, Itakura's monkeys continued to fail tests for self-recognition.

Given that animals can possess the capacities for locating mirrored objects and mirror-guided movement whilst still failing to recognise themselves, these other behaviours must not be reliant on recognising one's own reflection. Any account of self-recognition must consequently isolate the difference between these other capacities and self-recognition, so as to explain how monkeys can use mirrors in these ways whilst failing to self-recognise. Plausibly, what this means is that self-recognition cannot simply be the result of having some basic knowledge of how mirrors work, or the ability to use displaced visual feedback, as Celia Heyes (1994) suggests. This is the same capacity used in cursor tracking tasks, which monkeys can perform (Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989). Whilst capacities like this, which don't require any identification of the reflection with oneself, might underlie mirror-guided movement and locating mirrored objects, self-recognition evidently makes additional demands (Mitchell, 1993).

3. The Psychological Self-Concept Account

Gallup argues that passing the mark test demonstrates possession of a self-concept and a theory of mind. His discussions of the subject often begin by stating that mirror self-recognition requires self-awareness, where this is defined as the 'ability to become the object of one's own attention' (Anderson & Gallup, 2011; Gallup, Anderson, & Shillito, 2002; Gallup, 1985, 1998). It should be noted that 'x is the object of one's attention' can be read transparently or opaquely – since one can attend to an object that is in fact oneself without realising that it is oneself. Gallup's definition of 'self-awareness' requires the opaque reading (2002, p. 329) – so a subject is self-aware on this account if it is able to attend to an object that it identifies as itself. Gallup (1977, p. 334) writes that this kind of self-awareness is to be distinguished from any consciousness an organism has of its features by way of visual, tactile, chemical and proprioceptive feedback.

This starting point seems rather innocuous. On the definition of self-awareness offered, it seems almost analytic that self-awareness is what self-recognition requires – self-recognition just is a matter of identifying an object as oneself. That there is a distinction between this capacity and various forms of sensory awareness of oneself should also be uncontroversial, since all animals presumably have sensory awareness of themselves in some form, and yet only a fraction of animals recognise their reflections.

Gallup's development of these ideas, however, is problematic. He writes that self-awareness differs from self-specifying sensory information because the former involves an 'identity' which the others lack. This 'identity' is one component of a 'well-integrated

'simulation' of self-recognition. Outside the primate order, there is some evidence of self-recognition in dolphins (Reiss & Marino, 2001), Asian elephants (Plotnik, de Waal, & Reiss, 2006) and magpies (Prior, Schwarz, & Güntürkün, 2008).

self-concept', which has two additional components: 'a sense of psychological continuity over time and space' and 'an important sense of personal agency' (1998, p. 230). The sense of identity arises through social interaction, leading the subject to realise that it is 'both similar to and at the same time different from others'. This suggests that Gallup's self-awareness involves a rich idea of oneself in terms of a complex of properties, particularly psychological properties, which serve to distinguish one from other agents.

This sense of identity, although it grows out of social interaction, in turn provides the springboard for an understanding of the minds of others: 'if you are self-aware, then you are in a position to use your experience to model comparable experiences in others […] Knowledge of self is an inductive springboard for an inferential knowledge of others' (Gallup et al., 2002, p. 329). It should be highlighted here that if the sense of identity is to put one in a position to reason about the minds of others by analogy with one's own, then it must include substantial information about one's own mind and mental states, for instance, the circumstances that cause one to enter particular mental states, and the behaviour those states cause in turn.

Gallup has been extensively criticised for assuming that self-awareness is an all or nothing phenomenon, failing to consider alternative explanations, and not providing substantive arguments for his interpretation (see, for instance, Heyes, 1994; Mitchell, 1993, 1997; Morin, 2011; Savanah, 2013). Whilst a self-representation of some kind must be involved in self-recognition, this need not be nearly as complex as Gallup claims.

Language gives us a basis for thinking that there may be more and less complex kinds of self-representation. In language, we can distinguish between referring to oneself reflexively, using the first-person pronoun, and referring to oneself descriptively, for instance as 'the person writing this paper'. This distinction is similar in some ways to that between descriptions and names. Generally, when one uses either a name or 'I', one does so in the context of predicating some property of the referent – say, 'I'm hungry' or 'Fiona is hungry'. Each sentence predicates hunger of an object; importantly, neither embodies a commitment to any further claim about the kind of thing the object is. This is the sense in which the object terms are not descriptive. Where they differ, of course, is that 'I' is an irreducibly first personal representation. This means that it can't be given a third personal paraphrase whilst preserving its significance for the thinker, where this significance amounts to its immediate consequences for the thinker's actions (that is, knowing that I'm hungry will motivate me in ways that knowing Fiona's hungry will not, even if I am Fiona) (Perry, 1979).

The contrast between these two kinds of self-representation has been given in terms of language – but nothing hangs on this. When we talk about the thoughts of non-linguistic creatures, it is inevitable that we use a linguistic gloss to describe them. To suppose that this distinction between forms of self-representation exists in animal thought is just to suppose that there are some thoughts we would linguistically gloss in the one way, and some we would gloss in the other. If this is right, then from the mere fact that self-recognition involves self-representation, it does not follow that it involves a descriptive representation of oneself.

But even if self-recognition does require descriptive self-representation, the descriptions involved can be more or less complex. The complexity of descriptions can vary depending upon the concepts they employ. So, instead of descriptively representing myself as a 'person', I might more simply represent myself as a 'creature' or a 'thing'. The self-concept Gallup describes clearly involves highly complex concepts – yet it is not obvious that mirror self-recognition should require such a complex descriptive self-representation.

Gallup claims, for example, that the self-concept involved in self-recognition requires a sense of psychological continuity over time. This is implausible – there is no

obvious sense in which a mirror task singles out any psychological properties of the subject (Mitchell, 1997). It's highly likely that the mark test requires a psychological act - making use of one's kinaesthetic and visual capacities in a certain way, in order to 'match' their respective information. There's simply no reason, however, to suppose that it involves thinking about these capacities. Even if it did, it would not involve thinking about any other psychological capacities, and so there would be no reason for supposing that an animal passing the mark test possessed any other psychological concepts. As such, its ability to think about its thoughts would be far too limited to provide a conception of oneself as psychologically continuous – much less an 'inductive springboard' for mindreading, as Gallup suggests.

In a pair of recent papers, Savanah (Savanah, 2012, 2013) attempts to rehabilitate the psychological self-concept view of self-recognition by arguing that passing the mark test provides indirect evidence of a concept of self-as-agent – that is, 'an understanding of one's own existence as a psychological subject with intentional agency' (2013, p. 659). His argument is that passing the mark test demonstrates 'symbol-mindedness', which in turn is evidence of concept possession – in his view, a necessary and sufficient condition for the possession of a concept of self-as-agent. But this defence of the self-concept view is problematic in several respects.

First, the claim that symbol-mindedness is evidence of concept possession is unconvincing. Savanah paraphrases having conceptual capacities as 'the ability to ascribe meaning to things' (2013, p. 666), and claims that taking something to be a symbol also requires ascribing a meaning to it. Consequently, it seems an easy step to the claim that symbol-mindedness requires conceptual capacities. But closer examination suggests that 'ascribing a meaning' does not have the same sense in both contexts. Savanah writes that to have a concept of something is just to 'understand what type of thing it is' (2013, p. 666) – that is, to categorise it. Conceptual understanding is holistic – one cannot have the concept of 'fire', say, without the concepts 'hot' and 'burn' (2012, p. 716). Using concepts to 'ascribe a meaning' to something means categorising it – subsuming it under a concept. By contrast, treating something as a symbol according to Savanah means ascribing a meaning to it in another way – one treats it as 'standing in for' something else (2013, p. 666). A painting can have a meaning ascribed to it in both ways. One can categorise it as a painting, and perhaps relate this to one's concepts 'artist', 'paint' and so on; and one can treat it as 'standing in for' whatever it represents. But these are clearly two different activities: one doesn't have to do any categorisation to treat the painting as a symbol; and one needn't treat it as a symbol in order to categorise it. The connection Savanah draws between symbolic and conceptual thought rests on an equivocation.

Second, the claim that concept possession is a sufficient condition for a concept of self-as-agent is poorly motivated. Savanah argues that because concepts are holistically bound up with other concepts, they will all eventually be bound up with this self-concept. This seems optimistic. Whilst it may be right to understand concepts holistically – to think, for instance, that one couldn't grasp the concept 'fire' without the concept 'burn' – it seems possible that one could have a more or less complete grasp of a concept, and so that one's 'web of concepts' might be incomplete or fragmentary. One might for instance have some understanding of the concept 'blade' by relating it to 'cut', and perhaps 'action', but not know that actions are things performed by agents, which are psychological beings – much less know that one falls under that description oneself.[3]

It seems an obvious truth that self-recognition involves identifying an object as oneself. But there simply seems to be no good reason for thinking that this in turn involves psychological self-awareness – that is, knowledge of one's psychological

---

[3] For further discussion of these issues see Glock (2000) and Beer (1997).

properties. In the next section, therefore, I turn to deflationary accounts, which do not posit psychological self-awareness as an essential element of self-recognition.

4. Deflationary Accounts

In this section, I outline two deflationary accounts – the kinaesthetic-visual matching model, and the secondary representation account. Both deny that psychological self-awareness is required for self-recognition. Each instead argues that what is required is knowledge of one's physical appearance – a form of bodily self-awareness. This knowledge partially constitutes the evidence on which an animal is led to identify itself with its reflection. This seems absolutely right. However, I argue that these deflationary accounts do not fully explain why monkeys, which can use mirrors in certain ways, fail to self-recognise. In the final section, I argue that this is because identifying oneself with an object requires a form of self-awareness I call 'objective self-awareness', which is distinct from bodily self-awareness.

The kinaesthetic-visual matching account has been defended at length by Mitchell (1992, 1993, 1997, 2002, 2013), who takes self-recognition to rely on two capacities: understanding mirror correspondence, and kinaesthetic-visual matching. Understanding mirror correspondence amounts to recognising that 'mirror images (other than its own) are contingent accurate images of things outside the mirror' (1992, p. 129). In this definition, Mitchell inserts the caveat 'other than its own' because many animals are able to use mirrors for the purpose of locating objects and guiding their movements despite failing the mark test and not displaying any self-directed behaviour (as noted in §2). Mitchell's definition is meant to allow for these animals to understand mirror correspondence despite not recognising their reflections. For this reason, understanding mirror correspondence is by itself insufficient for self-recognition.

Kinaesthetic-visual matching is the ability to form a visual mental image of one's body on the basis of kinaesthetic feedback. This image is dynamic, in the sense of showing how the body's appearance changes when it moves, and is available for comparison with visual stimuli. Mitchell (2002, p. 364) describes this image as 'a general (and imprecise) idea of the "outline" of our bodies, and the relative positions of each part', suggesting that this image may be little more than a silhouette which is silent on the more fine-grained details of one's appearance such as colour, fur coverage and so on. This mental image can be compared with the reflection, and judged according to similarity with it. By comparing the two, the organism comes to realise that the reflection resembles the organism's body, and that its movements are contingent with the organism's. Because it knows that reflections are images of the things they resemble and move contingently with, it can therefore infer that the mirror image is 'of its body' (1993, p. 300). On this account, then, self-recognition is a simpler matter than Gallup takes it to be. It requires self-awareness, but only in the sense of an 'implicitly present' representation of one's body, which 'need not be reflected upon' (1993, p. 313).

A second deflationary account, initially suggested by Josef Perner (1991), and defended more recently by several others (Bard, Todd, Bernier, Love, & Leavens, 2006; Suddendorf & Butler, 2013; Suddendorf & Whiten, 2001), draws on the notion of a 'secondary representation'. In contrast to primary representations, which faithfully model reality, a secondary representation can be 'decoupled from reality' to represent past, future or hypothetical situations. Perner suggests that 'to understand that an image in the mirror is himself, the child needs one model [primary representation] to represent himself in reality and another [secondary representation] to represent himself in the mirror.'

An important difference between these deflationary accounts is that whilst the kinaesthetic-visual matching model proposes that knowledge of one's physical appearance derives from a specific capacity, kinaesthetic-visual matching, defences of the secondary representation account are often less clear about where this knowledge must come from, or what form it must take (Bard et al., 2006; Perner, 1991; Suddendorf & Whiten, 2001). It is consistent with these accounts that it derives from kinaesthesia – in fact, kinaesthetic-visual matching might even be viewed as an instance of secondary representation, since it enables subjects to compare multiple representations of the same thing.

However, Suddendorf and Butler (2013) also propose that the primary representation of oneself might instead derive from visual perception of one's physical features. This claim is motivated by the results of a study showing that infants perform worse at self-recognition tasks when their clothing is surreptitiously altered before the mark is applied. When they were allowed to update their knowledge of what clothes they were wearing and tested again, they performed much better (Nielsen, Suddendorf, & Slaughter, 2006). Suddendorf and Butler (2013) take this to show that visual knowledge of one's appearance can be an important factor in self-recognition, in providing the primary representation with which the mirror image is compared.[4] This seems independently plausible, since the mark test is premised on the existence of expectations about one's visual appearance. Unless the mark is contrary to the subject's expectations, the subject will have no interest in it. Consequently, it may be an advantage of the secondary representation account that it is agnostic about the origin of the primary self-representation – since this allows both visual and kinaesthetic knowledge of one's appearance to play a role in self-recognition.

Although they are distinct, in broad outline, these deflationary accounts have a common structure. Both claim that in order to self-recognise, one must have knowledge of one's physical appearance. The secondary representation account is agnostic about the origin of this knowledge – it could derive from vision, kinaesthesia, or perhaps even from something else. The kinaesthetic-visual matching account claims that this knowledge comes from a specific mechanism, kinaesthetic-visual matching, which derives an image of one's outline from kinaesthetic feedback. This knowledge of one's appearance is compared with the mirror image. In combination with some further capacity, the result is self-recognition. According to the secondary representation account, this relies upon the capacity for secondary representation, which enables one to compare and collate multiple representations of a single thing. According to the kinaesthetic-visual matching model, the kinaesthetically derived image is already directly comparable with the mirror image, enabling subjects to realise that the reflection *looks like me*. The realisation that it *is* me relies on understanding mirror correspondence, along with a general capacity for inference.

Both accounts demonstrate that self-recognition can be achieved without psychological self-awareness – that is, knowledge of one's psychological properties. In both cases, the suggestion is that self-recognition requires knowledge of one's bodily appearance – bodily self-awareness – along with some more general capacity for inference or comparing representations. As such, both are more plausible than the psychological self-concept view. However, both are incomplete in a way that leaves them unable to account for the behaviour of animals like monkeys, which can use mirrors in certain ways but cannot self-recognise.

According to the kinaesthetic-visual matching account, self-recognition is an inference drawn on the basis of understanding mirror correspondence and having

---

[4] But see Mitchell (2010) for critical discussion of the Nielsen et al. study.

knowledge of one's body shape provided by kinaesthetic-visual matching. The problem is in explaining why kinaesthetic-visual matching should be required – why is the same inference not made possible by the detailed information about one's appearance provided by visual self-perception? Monkeys have knowledge of their appearance from visual self-perception – and they have some understanding of mirror correspondence. As such, it is unclear why they do not self-recognise. It is no explanation to say that monkeys understand that mirror images 'other than their own' are accurate images of things in the environment. Apart from anything else, it is unclear what this claim amounts to – monkeys surely do not think something like 'all mirror images with the exception of my own are accurate images of things outside the mirror'. Of course something prevents monkeys from generalising their understanding of mirrors images to the case of their own reflections. But it would be incredible to suppose that they explicitly exclude their own reflection from a general belief they have about reflections.

The critical factor here is supposed to be that monkeys lack kinaesthetic-visual matching. The problem with this is that kinaesthetic-visual matching seems simply to be one source of knowledge about one's appearance. But information about one's appearance is also available from vision – and if the information provided by visual self-perception is not sufficient for drawing the inference that the reflected image is me, there is no immediate reason to think that the information about one's appearance provided by kinaesthetic-visual matching should support the inference either. This presents a dilemma: either there is something special about the information provided by kinaesthetic-visual matching which distinguishes it from that provided by visual self-perception, in which case more needs to be said about what that is, or the lack of kinaesthetic-visual matching does not fully explain monkeys' failure to self-recognise.

A related problem arises for the secondary representation account. On that view, self-recognition is merely one application of a more general ability to compare multiple representations of the same thing. This is appealed to by Suddendorf and Butler (2013) to explain individual variations in self-recognition. The suggestion appears to be that whilst the more general trait is uniformly present in the great apes, this particular application of it may not be uniformly distributed. This raises a question: what is special about *this* application of secondary representation? The question is again particularly acute when we reflect upon the various competences monkeys display with mirrors. In locating mirrored objects and in mirror-guided movement, monkeys appear to be using the mirror to form multiple images of objects in their environment – just not of themselves. Yet monkeys must have some visual knowledge of their appearance. So why should they have any special difficulty in using the mirror to entertain and compare multiple representations of themselves?

In both cases, it seems that the resources available to monkeys *should* enable them to recognise themselves in the mirror, just as they can recognise other objects. In the next section, I argue that the problem arises because there is something distinctively challenging about the final step in both accounts – drawing the inference 'that's me', or matching two representations of oneself. To see this requires us to return to the obvious observation with which we began – that self-recognition requires identifying an object as oneself. I argue in the next section that this requires a form of self-awareness I call 'objective self-awareness', which is distinct from bodily self-awareness. If a creature lacked this form of self-awareness, it would be blocked from taking the final step in self-recognition. I further argue that it is because the deflationary accounts discussed here either omit this fact or fail to make it explicit that they are unable to explain monkeys' failure to self-recognise. As such, whilst it is true as both accounts claim that bodily self-awareness is required for self-recognition, objective self-recognition is an additional requirement.

5. Self-Identification

Deflationary accounts are very likely correct in claiming that self-recognition involves drawing an inference from knowledge of one's appearance to the conclusion that 'that's me'. However, because neither discusses the importance of self-identification, both are incomplete. To address this problem, in this section I discuss what it means to identify an object as oneself. To do this, one must entertain a thought whose natural linguistic expression would be something like, 'that's me', or 'I'm the thing in the mirror'.[5] The problem presented by self-recognition might simply be that this is a complex or difficult thought to entertain, and consequently that the capacity to entertain such a thought is present in relatively few organisms.

Clearly, in order to have a thought like 'that's me', one must have the capacity to entertain first-person thoughts – thoughts whose natural linguistic expression makes use of the first-person pronoun. But it has been argued that there is a kind of first-person thought which is widespread in the animal kingdom (Bermudez, 1998). Like all first-person thought, it has irreducibly first-personal significance for the thinker, and so is intrinsically motivational (Perry, 1979). In addition, it has the property of being 'immune to error through misidentification relative to the first person pronoun' (Evans, 1982).[6]

First-person thoughts with this property are 'identification free'. This means that they cannot be decomposed into a predication component and an identification component. For instance, if the thought 'I'm angry' is immune to error through misidentification, then the person thinking it has not thought 'someone is angry', and then identified herself as that someone. Instead, the property of being angry is predicated immediately of the thinker, with no identification required. As such, the thinker cannot sensibly wonder whether it is she that has the property.

Immunity to error through misidentification is not a property of all first-person thoughts, nor even of particular first-person thought *types,* like 'I'm angry'. This is because immunity to error through misidentification is a property that arises only when a judgment is made on the basis of the right kind of information (Evans, 1982). If I judge 'I am angry' on the basis of some immediate sensory or introspective awareness of my anger, then the thought will be immune to error through misidentification. However, I might come to judge 'I am angry' in a different way – for instance, seeing a live video feed of a group of people, I might note that one of them has a very angry expression, and think 'someone is angry'. If I then come to believe that I am the angry-looking person, I will judge 'I am angry'. Because this judgment relies on identifying myself with the angry person, it is vulnerable to error through misidentification. (Of course, in both cases, the judgment can be mistaken – I might not, in fact, be angry.)

Whether it's true that animals have first-person thoughts of this kind (which are immune to error through misidentification) consequently amounts to the question whether they have representational states which are intrinsically motivational and are made on the basis of the right kind of information. The right kind of information is the

---

[5] Here, 'that's me' is meant to convey more or less what an adult human would mean by saying this about their reflection, or a photo etc. As such, thinking this thought does not involve thinking I'm numerically identical with an object 'over there', or that I'm multiply located – but it plausibly does mean thinking more than 'that looks like me'. 'That' is simply shorthand for an appropriate third-personal representation of oneself – as will become clear later in this section. Thanks to an anonymous reviewer for pressing this point.

[6] Hereafter simply 'immunity to error through misidentification'.

sort that, when it provides evidence that something is F, doesn't leave open the question of who or what is F. It has been argued that kinaesthesia and proprioception provide the right kind of information for judgments which are immune to error through misidentification (Bermudez 1998). If I come to believe, for instance, 'my arm is moving' on the basis of kinaesthetic awareness, this belief does not depend upon a belief like 'I am the subject of this instance of kinaesthetic awareness'. There is just no question of being moved to judge on the basis of that kinaesthetic information that someone else's arm is moving. Similarly, sensations like hunger, thirst and pain are intrinsically motivational, and do not rely on the subject believing anything like 'I am the thing which is hungry'. The question of who is hungry, thirsty or in pain is not left open.

Some of the information delivered by other spatial sensory modalities, in particular vision and audition, is 'self-specifying', and so can give rise to thoughts that are immune to error through misidentification. This is partly because they employ an egocentric frame of reference – so they implicitly provide an informational basis for judgments like 'there's a tree near me' which do not rely on any judgment of an identity between me and that object near the tree. Again, the visual information giving rise to this judgment doesn't leave open the possibility that it's really someone else near the tree (Evans, 1982). In addition, Bermudez (1998) argues that direct visual perception of one's body parts is immediately self-specifying in the same way, because one's body forms part of the invariant, egocentric structure of the visual field. If this is right, direct visual access to aspects of one's physical appearance can also give rise to this basic kind of first-person thought.

In general, information sources that give rise to first-person thought of the relevant kind have one of the following properties. They either provide information only about the self (as in kinaesthesia and proprioception), or the information they provide is given in an egocentric frame of reference, in which the self does not explicitly appear (as in vision and audition). In both cases, the self does not need to be explicitly *identified* as the subject matter of the information – the information is simply known immediately to be about the self. Given the numerous sources of what we might call 'first-personal information' that are available to animals, then, it seems appropriate to say that they have this kind of first-person thought. Indeed, this is a basic survival necessity: organisms need to be able to recognise information relating specifically to themselves, distinguish it from other information and act appropriately. I call the capacity for this most basic kind of first-person thought, which is intrinsically motivational and immune to error through misidentification, 'subjective self-awareness'.

Subjective self-awareness can be either bodily or psychological, but it is not built into its nature that it is either. In subjective self-awareness, one simply predicates some property of oneself on the basis of some immediate evidence that one has that property. Given the diversity of information sources which produce judgments which are immune to error through misidentification, the property in question could be either psychological or bodily. From introspection or sensory awareness, one might realise that one is angry, hungry or tired; from proprioception or visual self-perception, one might recognise that one is slouching, moving a particular way, or hairy. Whether one's thought manifests psychological or bodily self-awareness will just depend on what kind of property one happens to be representing.

Importantly, nothing in subjective self-awareness requires having a conception of oneself as being one kind of thing rather than another. As outlined in §3, in language we can refer to ourselves either reflexively, using 'I' or 'me', or descriptively. A reflexive self-representation, like a name, does not have any descriptive content. So by representing oneself in this way, one does not express any commitment about the kind of thing one is. Given what has been said about the thoughts involved in subjective self-awareness, it

seems plausible that their most natural linguistic gloss involves 'I', rather than anything more descriptive. The sources of subjective self-awareness provide information that is immediately taken to be about the subject. But there is no reason to suppose that, in providing this information, they furnish a subject with any descriptive information about the kind of thing it is. The information might represent the subject as having a certain property at a certain time (being angry, hungry, slouching, etc.) – but these properties need not be attached to a descriptive self-representation. It might be that a descriptive self-representation might eventually be built up out of such self-ascriptions of properties – but not necessarily so. This, presumably, will depend upon the memory and other cognitive capacities of the subject. Given, then, that the self-representation involved in subjective self-awareness does not have any descriptive content, and that it can be conjoined with representations of either bodily or psychological properties, it seems right to say that it is not constitutively a form of either bodily or psychological self-awareness.

Of course, to succeed in the mark test, one need only represent bodily properties. But doing this in the context of subjective self-awareness will be insufficient for self-recognition, for two reasons. First, subjective self-awareness is far more widespread than the capacity for self-recognition, since it simply falls out of having access to information sources like vision and proprioception. Second, subjective self-awareness is defined by not involving self-identification, upon which the capacity for self-recognition constitutively relies.

To see this, we can compare the situation of an animal looking in the mirror for the first time to a problem faced by Winnie the Pooh when he follows some tracks in the snow around a thicket, believing that they've been left by a 'woozle'. Eventually, Christopher Robin tells him that he has been walking in circles, prompting Pooh finally to realise that he was the one leaving the tracks all along (Milne & Shephard, 2006). When Pooh learns this, he forms a first-person thought which is vulnerable to error through misidentification – because it not only depends upon, but is entirely constituted by an identity claim. He has a thought which takes the form of an identity statement, in which 'I'/'me' and 'the woozle' (or 'the thing leaving the tracks') are the relata. These are both, strictly speaking, representations of Pooh. However, one is the familiar reflexive self-representation used in subjective self-awareness; the other is a description or name which is third-personal, in the sense of lacking the distinctive properties of first-personal representations such as immediate cognitive significance, consequences for action or guaranteed self-reference.

Pooh's is not a case of self-recognition – it does not require recognising something which looks like him. But it does require self-identification – realising that something (the leaver of these footprints) *is* him. Recognising one's reflection for the first time is similar in that respect, because it requires the same kind of thought. The subject is confronted by its reflection; in order to respond appropriately, it must recognise that the thing it sees *is* itself. This differs from subjective self-awareness in requiring the subject to represent an identity relation in which it stands as a relatum – and hence, to entertain a thought which is vulnerable to error through misidentification. As such, the thought that self-recognition depends upon an 'identity' which is absent in proprioception and kinaesthesia is correct. But the identity in question is merely the relation of identity, and not, as Gallup suggests, a complex self-representation making reference to a rich set of properties, including psychological properties, which distinguish one from other agents.

Precisely because they are vulnerable to error through misidentification, these self-identifications are of a different kind to the first-person thoughts involved in subjective self-awareness. Nevertheless, they're in a way parasitic on those thoughts. For 'that's me' to have its distinctive first personal significance – that is, for the subject to realise it's me in the mirror – it must feature the first-personal representation familiar from subjective

self-awareness. Without this, it would simply be a third-personal identification. Importantly, there is no tension involved in this claim. To claim that these thoughts share a common self-representation with thoughts that are immune to error through misidentification is not in any way to claim that they themselves are immune to error through misidentification. It is just to claim that they have the irreducible first-personal cognitive significance and motivational consequences that define first-personal representations. Immunity to error through misidentification is not a property of first-personal representations generally, but of first-personal thoughts formed on the basis of a particular kind of information.

A thought like 'that's me', when had about a reflection, relates a representation that's informed by subjective information sources to one that's informed by objective information sources. The problem with this is that the subjects of the representations will naturally seem entirely different from the subject's point of view. A reflexive self-representation like 'I' or 'me', although used in the context of predicating properties of oneself, does not itself have any descriptive content. In using it, one does not express any view about the type of thing one is. Moreover, it has irreducibly first-personal cognitive significance and consequences for action that cannot be preserved in any third-personal paraphrase. Of course, these are features of the representation, which may stem from the irreducibly first-personal information on which subjective self-awareness depends, in which the self is not identified as the subject matter. Thinking of oneself in this way, using a reflexive, non-descriptive and irreducibly first-personal representation, might have the effect of making the thing it picks out seem quite unlike other objects in the world.

As such, a subject could be forgiven for taking herself to belong to a different kind than other objects. This is not the result of metaphysical deliberation, but of making use of irreducibly first-personal information. Consequently, it will take a certain amount of cognitive sophistication to treat a thought like 'that's me' or 'I'm the thing in the mirror' as anything but a category mistake. Entertaining this thought requires thinking of oneself in a new way – not only as a subject, but as an object belonging to the same fundamental kind as other things in the world. This means extending the reflexive self-representation involved in subjective self-awareness beyond its usual use in first-personal predications based on immediately first-personal information, and conjoining it with some description of an object in the world – 'I am the thing in the mirror'. The extension of one's self-representation into this new context involves adopting a more objective perspective on oneself, and extends one's possible ways of learning about oneself. As such, it's a step to developing a more descriptive self-conception – an idea of the type of thing one is. We might call this awareness of oneself as identical with an object in the world 'objective self-awareness'.

That objective self-awareness is distinct from bodily self-awareness should now be clear. Like subjective self-awareness, it involves a reflexive self-representation with no descriptive content, which does not in itself appeal to any bodily or psychological properties. Objective self-awareness, however, involves conjoining this representation with an objective representation of an object. Again though, this representation need not appeal to any particular type of property. Consequently, the distinctions between psychological and bodily self-awareness and between objective and subjective self-awareness are independent and cross-cutting. Importantly, in the case of self-recognition, the evidence motivating the self-identification is likely to be information about one's bodily properties – so my claim is not at all that self-recognition does not require bodily self-awareness, but that it additionally requires objective self-awareness, and that these two forms of self-awareness are in principle independent of one another.

This distinction between subjective and objective self-awareness should not be confused with another distinction between forms of self-awareness – between the conceptual and the nonconceptual (Bermudez, 1998; Musholt, 2015). That distinction is between forms of self-awareness which do, and forms which do not, require mastery of a 'self-concept'. Precisely where this distinction is drawn depends upon one's account of concept possession. On one influential account, thought is conceptual when it meets Evans' (1982) Generality Constraint – that is, when its component representations can be generally and flexibly recombined to generate many more thoughts. For instance, a creature with the concepts of 'bone', 'tree', 'on' and 'under' might be able to think 'the bone is on the tree' or 'the bone is under the tree' or 'the tree is on the bone' or 'the tree is under the bone'. Conceptual thought, then, is characterised by its generalizability, systematicity and productivity. In addition to this, conceptual thought is often thought to be stimulus independent (Camp, 2009) - that is, conceptual thoughts can be entertained independently of any particular context.

Although all the examples of subjective self-awareness I have considered are, in this sense, nonconceptual, in claiming that self-recognition requires objective self-awareness I am not claiming that it involves conceptual self-awareness. Without considering further evidence, it would be difficult to claim that animals that can self-recognise are able to think objectively about themselves in a stimulus-independent way. Moreover, the self-representations of animals which can self-recognise may not be generally recombinable with other representations. Nevertheless, I would suggest that objective self-awareness involves a greater degree of flexibility in one's self-representation, since, as outlined above, it involves taking the self-representation normally used in the context of subjective self-awareness and combining it with a representation of a wholly different kind. But it is important that, in saying that self-recognition requires objective self-awareness, I do not intend to make a claim about concept possession. I mean only to claim that it involves representing oneself descriptively as an object in the world, rather than representing oneself only subjectively using a reflexive self-representation with no descriptive content.

I return now to the accounts discussed in the previous section. I argued there that both the kinaesthetic-visual matching account and the secondary representation account took self-recognition to be the result of comparing knowledge of one's own appearance with the mirror image, and inferring or recognising a match between them. I argued that, given that monkeys have knowledge of their own appearance and can recognise or infer matches between other objects and their reflections, this left it opaque why monkeys fail to self-recognise. The argument of this section is that drawing an inference like 'that's me', in which one matches two representations of oneself requires objective self-awareness.

This provides an answer to the question which arose in connection with the secondary representation account, namely, why self-recognition should present as a special instance of the capacity for secondary representation – why it should be different to having multiple representations of any other thing. The answer is that, in addition to requiring knowledge of one's physical appearance, it requires the capacity to match representations of radically different kinds. In other cases of matching multiple representation of the same thing, the representations are of the same general kind – they are third personal, objective representations informed by generic sources of information. In the case of self-recognition, one of the representations is like this, whilst the other is a reflexive self-representation informed by dedicated information sources, with no descriptive content. Matching these representations requires objective self-awareness – a recognition that one is an object like others in the world. This dependence of self-recognition on objective self-awareness renders the fact that self-recognition presents as

a special application of the more general capacity for secondary representation unsurprising. But it demonstrates that proponents of the secondary representation model are incorrect to claim that self-recognition requires self-awareness only in the sense of having knowledge of one's appearance (Suddendorf & Butler, 2013). A complete account of self-recognition must acknowledge that self-recognition requires objective self-awareness.

It also sheds light on the question which arose in connection with the kinaesthetic-visual matching account. I argued that it was unclear how kinaesthetic-visual matching could be the critical factor in monkeys' failure to self-recognise, since it simply provides more information about one's appearance. Yet having information about one's appearance is not sufficient for self-recognition, as the evidence from many animals attests. The argument of this section suggests that in addition to information about one's appearance, self-recognition requires objective self-awareness. Whether kinaesthetic-visual matching can be the critical factor in self-recognition, therefore, depends upon whether it involves objective self-awareness.

Mitchell (2003) writes that the 'self as an idea or object of thought' makes its appearance in kinaesthetic-visual matching – perhaps indicating a close connection between kinaesthetic-visual matching and objective self-awareness. But whether there is such a connection is a question which the characterisation of kinaesthetic-visual matching leaves open. Kinaesthetic-visual matching is characterised as the ability to form a visual mental image of one's outline on the basis of kinaesthetic feedback, which is available for comparison with other visual images. Although it involves an integration or 'matching' between kinaesthesia and vision, there's no suggestion that kinaesthetic-visual matching consists in a thought process, in which one identifies oneself with a third-personally represented object. Kinaesthetic-visual matching is simply the ability to derive a visual mental image of oneself from kinaesthesia, and compare it with other images. It is an open question at this point whether or not the information provided by this mental image is first-personal – that is, whether it is known by the subject to be a self-image.

One might think the mere fact that the information is presented in the form of a visual image seen as from the outside provides grounds for thinking that it is third-personal - and so not known to be a self-image. Although visual perception provides some self-specifying information, as discussed above, most objects presented in the visual field are not implicitly represented as 'me' - questions about their identity are left open. As such, a creature with such an image might take it simply to be an image it was controlling (as it might naively take its reflection) rather than an image of itself. In order to function as a self-image, the image would first have to be identified as such – and this would require having a thought like 'that's me' about it. Understanding kinaesthetic-visual matching in this way has the result that without the addition of objective self-awareness, organisms with kinaesthetic-visual matching would be unable even to recognise the similarity between themselves and their reflection – since they would not know that their mental image was a *self*-image.

This result demonstrates that this is not the best way to understand kinaesthetic-visual matching. It is clear that the explanatory role of kinaesthetic-visual matching is to explain how animals can recognise similarities between their own appearance and the appearance of other things. Consequently, it is more plausible to suppose that the image produced by kinaesthetic-visual matching is known by the subject to be a self-image. But now there is a further question: either the image is known immediately and without identification to be a self-image, or it is not, but kinaesthetic-visual matching partially consist in the ability to identify it as such.

The suggestion that it is known immediately and without identification to be a self-image fits well with Mitchell's (1993) claim that kinaesthetic-visual matching provides an

'implicitly present mental representation of the organism itself – a "self-representation" – which need not be reflected upon'. If it is right that the image produced by kinaesthetic-visual matching is subjective, organisms with kinaesthetic-visual matching would be able to recognise similarities between their appearance and that of other things. But this provides no reason for thinking that they would be capable of self-identification – i.e. realising not simply that I resemble the thing in the mirror, but that the thing in the mirror *is* me. If this is the right way to think about kinaesthetic-visual matching, therefore, it is not the (only) critical factor in self-recognition – objective self-awareness is an additional requirement.

The alternative is that the image is not known immediately to be a self-image – but that the ability to identify it as a self-image is constitutive of kinaesthetic-visual matching. This requires us to conceive of kinaesthetic-visual matching more broadly than Mitchell suggests – as encompassing not merely the ability to produce a kinaesthetically derived mental image of oneself, but also the ability to explicitly identify oneself with that image. Understood in this way, possessing kinaesthetic-visual matching would be sufficient for being objectively self-aware – since the capacity for self-identification would be an essential ingredient of kinaesthetic-visual matching.

The upshot for the kinaesthetic-visual matching account is not that it is false – but that it is incomplete in one of two ways. Either objective self-awareness is required in addition to kinaesthetic-visual matching for self-recognition to occur, or kinaesthetic-visual matching must be characterised more broadly, as encompassing self-identification and hence objective self-awareness. In either case, the role of objective self-awareness in the production of self-recognition remains to be explicitly articulated in this account. Only once the role of objective self-awareness is articulated can the account make sense of monkeys' failure to recognise their reflections.

The distinction between objective and subjective self-awareness is significant because it highlights the importance of the obvious truth that self-recognition requires identifying an object as oneself. Consequently, it shows that the deflationary accounts discussed here must make explicit appeal to objective self-awareness if they are to provide a complete account of self-recognition. More broadly, highlighting this distinction sheds important light on the understanding self-recognising animals have of themselves. If an animal recognises itself, this does not simply show that it has some understanding of what it looks like – it also shows that it is capable of adopting an objective perspective on itself, and so recognises that it is an element of the objective order. Consequently, investigations into self-recognition in animals have the potential to inform us about the distribution and origins of a kind of objectivity and detachment from one's immediate, subjective perspective on oneself.

Moreover, whilst we need not go so far as to say that self-recognition is itself evidence of a self-concept – a generally recombinable, flexible, stimulus-independent self-representation – it is a step in that direction. This is because it involves extending the reflexive self-representation belonging to subjective self-awareness and using it in an altogether different context to form thoughts of a different kind, which are vulnerable to error through misidentification. It seems plausible that this could be a stepping stone toward other cognitive capacities which require reflection on the self – including planning and 'mental time travel', both of which require thinking objectively about oneself in a stimulus-independent way. In addition, the ability to adopt an objective perspective on oneself might eventually enable one to conceive of the distinction between one's subjective experience and the way things actually are. This would be a crucial element of intentional deception and theory of mind, the litmus test for which requires animals to understand that others can have false beliefs. Mitchell (e.g. 1993, p. 313) makes a similar claim to the effect that the self-representation involved in

kinaesthetic-visual matching may be involved in some of these capacities – in particular, planned deception and pretence. Of course, if the claim is that kinaesthetic-visual matching involves thinking of oneself objectively, and these capacities also do, then our claims are quite compatible. Whether these capacities rely on having a particular kind of bodily knowledge, such as that discussed by Mitchell, is a separate question, and one I do not address here.

All this is not to say that evidence of self-recognition is itself evidence of these other capacities – but it seems likely that they require thinking objectively about oneself, and so that self-identification would count as an important step towards these more sophisticated abilities.

6. Conclusion

Recent debate about self-recognition has centred on whether it is evidence of psychological or only bodily self-awareness. The distinction between these two is simply the difference between ascribing psychological or bodily properties to oneself. I have argued that this question has been conclusively settled – self-recognition requires the ascription only of bodily properties to oneself.

However, framing the question in this way obscures the answer to the broader question: of what kinds of self-awareness is self-recognition evidence? This is because, as well as dividing into the bodily and the psychological, self-awareness divides into the subjective and the objective. Subjective self-awareness is the capacity for first-person thought which is immune to error through misidentification. This capacity, I have argued, is shared by many more creatures than the capacity for self-recognition. This asymmetry is due to the fact that self-recognition constitutively relies upon objective self-awareness – the capacity to entertain thoughts like 'that's me' or 'I'm the animal in the mirror', which are first-personal, but involve self-identification, and so are vulnerable to error through misidentification. These distinctions, between bodily and psychological self-awareness on the one hand, and subjective and objective self-awareness on the other, are cross-cutting. Both subjective and objective self-awareness can involve the ascription of either psychological or bodily properties. As such, the claim that self-recognising animals are objectively self-aware is distinct from the claim that they have bodily self-awareness – and both claims are required in a complete account of self-recognition.

The role of objective self-awareness in self-recognition raises an interesting possibility about variations in self-recognition both within and between great ape species. Self-recognition is a complex phenomenon, and so there are many possible explanations of this variation. One possibility indicated by this paper is that it results from variations in objective self-awareness – which may not be uniformly distributed across the great ape species. That there is cross-cultural variability in self-recognition in children (Kärtner, Keller, Chaudhary, & Yovsi, 2012), and that encultured gorillas are far more likely to self-recognise than their wild counterparts (Anderson & Gallup, 2015) suggests that social factors may have a role to play in the development of objective self-awareness. Since cultural variation in children is thought to be linked to the extent to which caretakers emphasise the child's individuality and autonomy (Kärtner et al., 2012), we might speculate that focussing on one's individuality facilitates thinking of oneself as an object – and perhaps extensive interaction with humans encourages gorillas to think of themselves as individuals in a non-species-typical way. Similarly, if the contested evidence of self-recognition in macaques (Chang, Fang, Zhang, Poo, & Gong, 2015) is borne out

by further studies this might suggest that objective self-awareness can develop with the right kind of training.[7]

Although self-recognition is evidence of bodily self-awareness, its real interest is that it is evidence of a kind of objectivity of thought – a realisation that one is an *object*, and so of the same basic kind as the things in one's environment. Realising this requires one to have a more flexible, and generalizable representation of oneself. This, though not necessarily a self-concept, is a necessary step toward more sophisticated cognitive capacities often thought to be distinctive of human thought, including planning, mental time travel and theory of mind.[8]

References

Anderson, J. R. (1986). Mirror-mediated finding of hidden food by monkeys. *Journal of Comparative Psychology*, *100*(3), 237–242.

Anderson, J. R., & Gallup, G. G. (2011). Which `primates` recognize themselves in mirrors? *PLoS Biology*, *9*(3), e1001024.

Anderson, J. R., & Gallup, G. G. (2015). Mirror self-recognition: a review and critique of attempts to promote and engineer self-recognition in primates. *Primates*, *56*(4), 317–326. doi:10.1007/s10329-015-0488-9

Bard, K. A., Todd, B. K., Bernier, C., Love, J., & Leavens, D. A. (2006). Self-awareness in human and chimpanzee infants: what is measured and what is meant by the mark and mirror test?, *9*(2), 191–219. doi:10.1207/s15327078in0902_6

Beer, C. (1997). Expressions of mind in animal behaviour. In R. W. Mitchell, N. S. Thompson, & H. K. Miles (Eds.), *Anthropomorphism, anecdote, and animals* (pp. 198–209). Albany, NY: SUNY Press.

Bermudez, J. L. (1998). *The Paradox of Self-Consciousness*. Cambridge, Massachusetts: MIT Press.

Camp, E. (2009). Putting thoughts to work: concepts, systematicity and stimulus independence. *Philosophy and Phenomenological Research*, *78*(2), 275–311.

Chang, L., Fang, Q., Zhang, S., Poo, M. M., & Gong, N. (2015). Mirror-induced self-directed behaviors in rhesus monkeys after visual-somatosensory training. *Current Biology*, *25*(2), 212–217. doi:10.1016/j.cub.2014.11.016

Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.

Gallup, G. G. (1977). Self-recognition in primates: a comparative approach to the bidirectional properties of consciousness. *American Psychologist*, *32*(5), 329–338.

---

[7] Thanks to an anonymous reviewer for pressing this point.

Gallup, G. G. (1985). Do minds exist in species other than our own? *Neuroscience and Biobehavioral Reviews*, *9*(4), 631–41.

Gallup, G. G. (1998). Self-awareness and the evolution of social intelligence. *Behavioural Processes*, *42*, 239–47.

Gallup, G. G., Anderson, J. R., & Platek, S. J. (2011). Self-recognition. In S. Gallagher (Ed.), *The Oxford Handbook of The Self* (pp. 80–110). Oxford: Oxford University Press.

Gallup, G. G., Anderson, J. R., & Shillito, D. J. (2002). The mirror test. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The Cognitive Animal* (pp. 325–334). Cambridge, Massachusetts: The MIT Press.

Glock, H.-J. (2000). Animals, thoughts and concepts. *Synthese*, *123*(1), 35–64.

Heyes, C. M. (1994). Reflections on self-recognition in primates. *Animal Behaviour*, *47*(4), 909–919.

Itakura, S. (1987a). Mirror guided behavior in Japanese monkeys (Macaca fuscata fuscata). *Primates*, *28*(2), 149–161.

Itakura, S. (1987b). Use of a mirror to direct their responses in Japanese monkeys (Macaca fuscata fuscata). *Primates*, *28*(3), 343–352.

Kärtner, J., Keller, H., Chaudhary, N., & Yovsi, R. D. (2012). The development of mirror self-recognition in different sociocultural contexts: Results. *Monographs of the Society for Research in Child Development*, *77*(4), 37–65. doi:10.1111/j.1540-5834.2012.00688.x

Menzel, E. W., Savage-Rumbaugh, E. S., & Lawson, J. (1985). Chimpanzee (Pan troglodytes) spatial problem solving with the use of mirrors and televised equivalents of mirrors. *Journal of Comparative Psychology*, *99*(2), 211–7.

Milne, A. A., & Shephard, E. H. (2006). *Winnie The Pooh*. London: Egmont.

Mitchell, R. W. (1992). Developing concepts in infancy: animals, self-perception and two theories of mirror self-recognition. *Psychological Inquiry*, *3*(2), 127–130.

Mitchell, R. W. (1993). Mental models of mirror-self-recognition: two theories. *New Ideas in Psychology*, *11*(3), 295–325.

Mitchell, R. W. (1997). Kinesthetic-visual matching and the self-concept as explanations of mirror-self-recognition. *Journal for the Theory of Social Behaviour*, *27*(1), 17–39.

Mitchell, R. W. (2002). Kinesthetic-visual matching, imitation and self-recognition. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The Cognitive Animal* (pp. 345–352). Cambridge, Massachusetts: The MIT Press.

Mitchell, R. W. (2010). Understanding the body as spatial cognition. In F. L. Dolins & R. W. Mitchell (Eds.), *Spatial perception, spatial cognition: Mapping the self and space* (pp. 341–364). Cambridge: Cambridge University Press.

Mitchell, R. W. (2012). Self-recognition in animals. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 656–679). New York: Guilford.

Mitchell, R. W. (2013). A critique of Stephane Savanah's "mirror self-recognition and symbol-mindedness." *Biology & Philosophy*. doi:10.1007/s10539-013-9403-1

Morin, A. (2011). Self-recognition, theory-of-mind, and self-awareness: What side are you on? *Laterality*, *16*(3), 367–383. doi:10.1080/13576501003702648

Musholt, K. (2015). *Thinking about Oneself.* Cambridge, Massachusetts: The MIT Press.

Nielsen, M., Suddendorf, T., & Slaughter, V. (2006). Mirror self-recognition beyond the face. *Child Development*, *77*(1), 176–185.

Perner, J. (1991). *Understanding the Representational Mind.* Cambridge, Massachusetts: MIT Press.

Perry, J. (1979). The problem of the essential indexical. *Nous*, *13*(1), 3–21.

Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(45), 17053–7.

Prior, H., Schwarz, A., & Güntürkün, O. (2008). Mirror-induced behavior in the magpie (Pica pica): evidence of self-recognition. *PLoS Biology*, *6*(8), e202.

Reiss, D., & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(10), 5937–42.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (Macaca mulatta), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology*, *103*, 32–38.

Savanah, S. (2012). The concept possession hypothesis of self-consciousness. *Consciousness and Cognition*, *21*(2), 713–720. doi:10.1016/j.concog.2011.02.019

Savanah, S. (2013). Mirror self-recognition and symbol-mindedness. *Biology & Philosophy*, *28*, 657–673. doi:10.1007/s10539-012-9318-2

Suddendorf, T., & Butler, D. L. (2013). The nature of visual self-recognition. *Trends in Cognitive Sciences*, *17*(3), 121–7.

Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: evidence for secondary representation in children, great ages, and other animals. *Psychological Bulletin*, *127*(5), 629–650. doi:10.1037/0033-2909.127.5.629