Origins of Self
Edwardes, Martin P. J.

Published by University College London

Edwardes, Martin P. J.
University College London, 2019.
Project MUSE. muse.jhu.edu/book/81330.

For additional information about this book
https://muse.jhu.edu/book/81330

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=2771971
'What do you mean by that?' said the Caterpillar sternly. ‘Explain yourself!’
'I can’t explain *myself*, I’m afraid, sir’ said Alice, ‘because I’m not myself, you see.’
'I don’t see,' said the Caterpillar.
'I'm afraid I can't put it more clearly,' Alice replied very politely, 'for I can't understand it myself to begin with; and being so many different sizes in a day is very confusing.'

(Lewis Carroll 1865, Chapter 5: ‘Advice from a Caterpillar’)

Poor Alice! Not being Alice means that she cannot explain Alice to the Caterpillar; even worse, she cannot understand why she cannot explain Alice, because the Alice she has been since falling down the rabbit-hole has been inconsistent and unanchored. The Caterpillar, on the other hand, has a very firm understanding of the self he is, and cannot understand why Alice is not equally certain about her self; yet, as Alice later points out, the Caterpillar will at some stage undergo his own transformation, much more fundamental than hers, when he chrysalises and transforms into a butterfly. Will the butterfly-self recognise the caterpillar-self? And, if so, will he remember the caterpillar-self as being himself, or just a self? Or will there be a major dislocation of memory and selfhood, so the butterfly does not even remember its life as a caterpillar? There is some evidence that moths can retain aversive behaviours learned as caterpillars (Blackiston et al. 2008), but it is unlikely that a moth or butterfly knows that it used to be a caterpillar. At least Alice has continuity in her selfhood, so that the Alice of today is a product of, and remembers, the Alice of yesterday – even if she cannot explain what it is that defines today’s Alice.
However, is there really a continuity of selfhood that humans can use to define their selves? Or is discontinuity actually the way our selves work? Is continuity yet another of the self-illusions (or self-delusions) that humans are so adept at modelling? To examine this debate, we need to look more closely at what kinds of self define us, how they define us and why we let them define us.

Let us start by setting out the types of self we need to talk about. Two of them have been discussed so far in some detail: the Social self and the Cultural self. Two have been mentioned in passing: the Episodic self and the Narrative self. And two have been rather underplayed, considering their importance: the self-model and the Projected self. That leaves the elephant in the room, which has been largely ignored so far: the actual, physical self that continues to sit on the sofa eating peanuts and ignoring all metaphysical attempts to wish it away. It is here that we need to begin our definition of selves.

To help us in our self-definition, we should start in a traditional fantasy-story way, with a map showing the journey ‘there and back again’. This does not mean that the story of the selves is a fantasy, but it is, like all good fantasy stories, a metaphor of a reality – and metaphors are both more entertaining and more revealing than gritty reality stories. Figure 8.1 shows us the map of our tour round the selves, and it also shows the relationships between the selves.

![Fig 8.1 Types of self](image)

The arrows indicate information flows. The big ones represent information flows into and out of the individual’s cognition, the smaller ones represent information flows within the individual’s cognition. The direction of the arrow is the direction of the flow.
The Actual self: unknowable

The fact that we have a physical self is inescapable: all selfhood resides in the brain, which is a component of the physical body; and it is the physical self in the brain that regulates the body. For practical purposes, and ignoring the unprovable non-physical, there is no existent self without the physical self. This makes the physical, or Actual, self a key feature of selfhood. The Actual self is a Darwinian gene-machine – unlike all the other selves, which are cognitive products of a Darwinian gene-machine – which means that the Actual self is directly governed by the twin genetic imperatives to survive and thrive. It has no interest in philosophical positions such as self-sacrifice or generosity, unless they directly lead to enhanced personal survival or enhanced reproductive success. So when William Hamilton said:

… in the world of our model organisms, whose behaviour is determined strictly by genotype, we expect to find that no one is prepared to sacrifice his life for any single person but that everyone will sacrifice it when he can thereby save more than two brothers, or four half-brothers, or eight first cousins …

(Hamilton 1964, 16)

… he was speaking for his Actual self. Three brothers, five half-brothers or nine first cousins all have more of my genes that I do – that is, if I can be assured of no sexual cheating by my mother, grandmother or aunts. In the actual world, this Hamiltonian kinship calculus is hedged-around with uncertainties, and not just in terms of non-monogamous relatives; so the safest strategy for the Actual self remains, look out for number one.

Unlike the other selves, the Actual self is not a conscious representation of the self. Yes, invasions of the self–other boundary cause discomforts or pleasure – injury causes pain, food causes satiation; but these are innate electrochemical reactions generated within the Actual self. They require no conscious attention – the Actual self does not need to be an aware self. An organism can handle pain and pleasure with its autonomic responses (involuntary, innate mechanisms) or its automatic responses (incidentally acquired mechanisms that are non-conscious but which may be subject to contextual override). For the organism to experience pain or pleasure, there is no need for a self that is able to recognise themself as being in pain or being happy.
Also unlike the other selves, the Actual self is not a constructed model. Where the other selves represent, the Actual self just is. And where the modelled selves are either differentiated (different selves of the same type) or integrated (the outcome of merging different self-models to create a new modelled self) or both, the Actual self is neither differentiated nor integrated. There is only one Actual self, and it is not the product of modelling or other cognitive manipulation. However, because it is not a conscious representation, it is also rather dull. It may be the elephant in the room, sitting on the couch and eating peanuts, but we have no need to disturb it. It is happy being ignored, although it is not necessarily aware of being happy.

The Social self: the self others believe me to be

The Social self is the first self of which we are consciously aware – the model of my self offered by others as part of the exchange of social calculus. Unlike most of the other selves, therefore, it is provided wholly from outside the self, and is therefore subject to the receiver’s dilemma that accompanies any shared information: why should I believe it? The answer lies in the peculiar nature of social calculus information. First, it is opinion, not knowledge, and it is offered and received as such; it already contains its own veracity warning. Second, it is offered about me to me: the first time this happens, I have no pre-existing data against which to check the offered model, but at all subsequent times I have a growing database modelling how I am seen by others, so I can accept or reject this new opinion based on that database. Third, if it is offered in a social calculus equation, then I can check the offered information against my own social calculus: does it mesh with the information already in my system or does it produce anomalies or contradictions? Fourth, any social calculus information offered is a World 3 reality, not a World 2 virtuality or a World 1 actuality; all information is subject to negotiation toward meaning – both with the sender and with my models of my self.

These caveats to veracity mean that all models of me offered to me are factually relative: their ‘truth’ is relative to the social calculus of the sender and to my own social calculus. I should not expect to receive social models of me that give a single group impression of who I am; instead, I will receive a number of different, individual views. Some of these views will differ from each other only slightly, while others will differ markedly; but their multiplicity means that the Social self is differentiated, with more than one model available to the self.
The self-model: the self I believe me to be

Accepting models of my self from others is informative in building a picture of how I am viewed by others; but, more importantly for the story of selfhood, it provides a third-person model of me to sit in my social calculus system – and this model appears to be undifferentiated. So how to collapse the many Social selves into one self-model? The obvious solution is to merge the different social models into an integrated self-model, discarding contradictions; but, because of the wide variation in Social selves offered to me, this is neither possible nor necessary. Instead, we tend to hold cognitive representations of several self-models, although only one at a time (the self-in-context) is treated as the valid self for modelling purposes. Over time, we cycle through a range of self-models depending on our current electrochemical state, the context in which we are self-modelling and the company we are keeping.

Each of my self-models is integrated from sets of social models, but because I have more than one self-model they are also differentiated; and because the currently active model changes over time, the self-model is also protean – like an amoeba, it constantly changes its form. This variable nature of the self can become an important source of self-anxiety: do my inconsistencies represent a failure or fraying of my integrated selfhood? There is only one actual, physical me, so why do there seem to be several real me’s?

One way out of the dilemma is to ignore the differences and simply believe there is a fully integrated self-model that collocates with the Actual self. This seems to be the solution chosen by narcissists, who usually avoid internalised self-awareness in favour of externalised approbation. Their internal self-models (their egos) are often brittle and easily damaged. Any criticism is not just criticism of the critic’s model of the narcissist, it is criticism of the single self-model the narcissist possesses; it is, therefore, criticism of the narcissist in every possible way. Critics are therefore dealt with vindictively and with an overkill out of proportion to the criticism offered (Campbell and Baumeister 2006). However, while this is a description of people who would be clinically diagnosed as narcissists, it is also a description of most of us at one time or another: we all sometimes believe that the currently active self-model is our only self-model, and take criticism badly. What is being described here is a strategy of selfhood, not necessarily an aberrant psychological behaviour (Krajco 2007). What makes it aberrant is if it is an individual’s default, or only, strategy.
A second way out of the dilemma is to believe that the differences amount to a negation of an integrated self-model. All I have to inform my self-model are the impressions that others have of me; and these can be manipulated by adopting whatever appearance-model will get me immediate gratification. This seems to be the solution of sociopaths (also called psychopaths), who take the protean aspect of self-modelling to extremes: they will do, say and be whatever they need to in order to manipulate others and satisfy their needs. There is no negotiation toward meaning for sociopaths because there is only one relevant meaning to the universe – that of the undifferentiated self. And criticism does not matter, because there is no self-model to criticise, only a projected model or appearance that can be adjusted to meet current needs. If someone does not like the self you are projecting, project a different self (Gallagher 2013b). To quote Groucho Marx, ‘Those are my principles, and if you don’t like them … well, I have others’.

Once again, it must be emphasised that what is described here is not a diagnostic aid for identifying clinical sociopathy. It is a self-modelling behaviour that we all display at one time or another. In fact, the frequency of clinical sociopathy in the general population (about 3 per cent of males and just under 1 per cent of females) makes them unusual but not rare (Mealey 1995). On any London tube train during rush hour, there are likely to be a dozen or more people who would be diagnosed as sociopathic if their behaviour warranted a diagnosis; but, in the vast majority of cases and for most of the time, it does not.

A third way out of the dilemma is not really a way out at all. It is possible to surrender to the multiple selves by losing the capacity to control which single self-model dominates at any one time. This is the problem that schizophrenics face, with the different self-models competing in the conscious mind, rather than being policed by the subconscious mind and presented to the conscious mind one at a time. In this case, my inconsistencies do not represent a failure or fraying of the integrated model of my selfhood, they really are the fraying of my selfhood. This collapsing selfhood causes the boundaries between actuality, reality and virtuality to blur even more than usual, hence typical schizophrenic symptoms include hallucinations and delusions. The lack of a cohesive self-model also affects the self that the person can project, causing breakdowns in their social relationships (Bowes 2014).

Schizophrenia is a particularly interesting ‘solution’ to the many-selves dilemma, because it seems to be a by-product of having language: the condition has been linked to language in several ways. First,
schizophrenia is linked to dysphasia, or the loss of communicative competency; and it also seems to affect phonology, leading to flat-toned speech (Covington et al. 2005). Second, schizophrenia has been shown to be implicated in the language and social-modelling areas of the brain. Radanovic et al. (2013) discovered a link between formal thought disorder (a diagnostic criterion for schizophrenia) and language impairment. The severity of both impairments was correlated with deficits in the left superior temporal gyrus and the left planum temporale, both areas in a Statistically Standard Brain (SSB) implicated in language; and in the orbitofrontal cortex, which is implicated in modelling for decision-making, including social calculus modelling. Pu et al. (2017) identified correlated deficits in the anterior part of the temporal cortex, the ventro-lateral prefrontal cortex, the dorso-lateral prefrontal cortex and frontopolar cortex areas of the brain – all areas implicated in both social cognition (particularly ToM) and language production. It seems that schizophrenia is somehow involved in the neural connections between social cognition and language cognition.

It also seems that there may be a genetic basis to the link between schizophrenia and language. The gene FOXP2 is implicated in both language production and hallucinatory episodes in schizophrenia (Tolosa et al. 2010); there does not seem to be a causal relationship, but there is a strong correlation. Srinivasan et al. (2016) showed that many of the genes implicated in schizophrenia are also involved in general cognitive development, and specifically human versions of the genes seem to have appeared since the split between humans and Neanderthals. These gene-forms seem to be absent from chimpanzee genomes (Srinivasan et al. 2017).

Dissociative Identity Disorder (DID) is another condition in which the individual seems to surrender to the dilemma of multiple selves. This condition has similarities to schizophrenia, and it is often presented in the lay media as schizophrenia. In terms of the SSMH, it poses a slightly different problem for the individual: it is not that the selves are indistinct, it is that there is no concord between them; each self has somehow set itself up as an independent person – sometimes with recognition that there is only one shared body, but sometimes not. Whitehead points out that we all live with multiple modelled personalities when he says, ‘Shakespeare can fill a stage with characters, all of whom act and speak convincingly as whole and distinct persons, though all were born within a ‘single’ mind’ (Whitehead 2001, 4). The problem with DID would seem to be that Shakespeare has left the stage, and all that is left are the characters.
Depersonalisation Disorder is yet another condition on the schizophrenia spectrum where SSMH may be relevant. In this condition, individuals feel they are somehow not a real person: the modelling is still working, but the self has gone missing (Baker et al. 2003). It is a state that we are all in at some point in our lives, and it is only a disorder when it becomes prolonged. This condition shows that the metaphor SELF IS OTHER is more than just a cognitive explanation after the fact; it is itself a cognitive fact with consequences.

Between too little modelling, too many selves, and not enough self, the schizophrenia spectrum seems to be a product of self-modelling. However, as with narcissism and sociopathy, it must be emphasised that what is being discussed here is a self-modelling behaviour, and does not necessarily indicate a medical condition. Poor Alice could not explain herself because a very confusing day had exhausted her range of self-models, leaving her with no dominant self to rely on; but she did recover quite spectacularly by the end of the book. In a similar way, we can be left nonplussed and dumbfounded without needing a diagnosis of formal thought disorder and dysphasia. Life may currently be complicated; this, too, shall pass.

Most of us wander around in this triangle of selfhood extremes without particular difficulty, doing what we need to get by as a self-modelling entity; and mostly we do it without being consciously aware of the choices we are making about our self-modelling. The terminology of selfhood modelling often deceptively implies conscious choices are being made in self-modelling; but most self-modelling involves subconscious cognition, implicit knowledge and automatic responses. We first start to build our self-models when we receive our first recognised piece of information about our self from others; and this happens around age 2, when the child becomes aware of the dyadic negotiation toward meaning between them and their caregiver. ‘Daddy loves Baby’ may sound like a simple idea for adults to comprehend, but it makes huge demands on the social calculus and social modelling of the child, long before the child has conscious knowledge of their social selfhood.

The Episodic self: the self as modelled in individual past events

The Episodic self is a feature that emerges from the combination of self-modelling and conscious memory recall. As an emergent feature, the Episodic self is neither directly learned nor directly innate; it is what becomes possible when there is an interaction between two other
features that are, themselves, learned or innate (Pomerantz and Cragin 2015). The emergent feature of the Episodic self seems to be particularly interesting, because one of the interacting features is learned and the other is innate: the combination of the capacity to self-model (learned) and the capacity to remember events in the past (innate) creates the possibility of modelling a self in a remembered past event. Instead of the event being passively visceral – the emotions of the event are remembered as emotions – it becomes actively visceral – the emotions of the event are remembered as my own emotions. The Episodic self is, therefore, more than just an episodic memory, and it is equally as real as any self-model.

An Episodic self is not a memory of a past self-model; it is a current representation constructed from the social-self evidence currently available. When we remember our self, we do not remember our self-model as it was when the memory was laid down; rather, we construct a current self-model to represent our previous self. Giorgio Marchetti (2014) says this is because we are prone to three ‘sins’ of memory: we forget or mitigate the visceral emotions that were actually generated by the event, making our emotional memory of the event unreliable; we distort our memories by remembering the events themselves incorrectly, thus rendering our procedural memory of the event unreliable; and we over-emphasise some aspects of the event while under-emphasising others, thus pathologising our memory of the event. To use a von Neumann computer metaphor, recalling a memory is not just accessing a fixed memory-image like a file on a hard drive; it involves copying the memory-image into working memory, adjusting the memory and then writing it back as a new image (Schiller and Phelps 2011). But memory is not just accidentally inherently fallible; it is important that it be so, so that each time I recall the memory I can model the experiences in the past event as my experiences in relation to my current self-model.

Yet another feature makes an episodic self-model unreliable as a model of my previous self: it is composed of more than my own memory of the event. The sharing of social models is more effective if the contexts and evidence of those models are also shared; so any memory I have of a past event is overlaid with the memories of others about that event – and every viewpoint of the event is different. What I believe is my memory of the event, seldom is; instead it is, like my models of my self and others, an amalgam of viewpoints and opinions. The Episodic self is a memory of a self-model that was originally generated from the models of me offered by others, and when recalled it is then edited by my current model of me and by more models of me offered by others.

If the Episodic self is just a type of self-model, and self-modelling is an outcome of shared social calculus, and sharing social calculus seems
to be limited to humans, then is it possible for non-humans to have Episodic selves? The simple answer would seem to be no; but, because the Episodic self is an emergent feature, the actual answer is more complex. One of the two components of an Episodic self is the capacity to recall past events, which Endel Tulving (2005) calls noetic memory, contrasting it with autonoetic memory (the capacity to recall past events that include the past self’s own perspective). Tulving takes the view that autonoetic memory is available only to humans. In his description, autonoetic memory seems similar to episodic memory, having both recall and a self-perspective. Tulving’s autonoetic self-perspective is that of a past self; but is that a true self-perspective, and is that past self really available to the current self as a self-model?

Some researchers disagree with Tulving’s view. Fabbro et al. (2015) propose that a capacity for autonoetic memory is likely to be present in non-human brains, because the neurologically complex brain areas associated with human selfhood have correlates in those non-human brains. However, it remains to be demonstrated that the correlate areas function in the same capacity in both brains. In contrast, Robert Numan (2015) differentiates non-human from human episodic memory by describing it as ‘episodic-like’. This is a more cautious approach that does not necessarily require the generation of an Episodic self. It also gives us a convenient way to label the difference between the autonoetic episodic selfhood of humans and the otherwise episodic memory that many mammals do appear to possess.

If episodic selfhood, like Tulving’s autonoetic memory, is an innate capacity in humans, then we have a problem explaining how it evolved. If, however, it emerges from the modelled self plus noetic memory, and the modelled self is an outcome of sharing social calculus, then we can say that episodic selfhood is a synthesis of pre-existing cognitive systems. It is a learned trick, a way of creating a third-person social calculus model that happens to represent the self. Autonoetic memory is not a species-difference requiring its own evolutionary explanation; it is just noetic memory plus a trick.

**The Narrative self: the remembered self, the self with history**

The concept of Narrative self, or narrative identity, was perhaps first codified by Paul Ricoeur (1990 [1992], Chapter 6) and Jerome Bruner (1990, Chapter 4). Although it had already been discussed by earlier
commentators, Ricoeur and Bruner were the first to define what a Narrative self is. Basically, the Narrative self is the model we make of our life experiences as an evolving story – a stitching-together of the various Episodic selves in such a way that they can be viewed as aspects of a single self. Where the Episodic selves, being self-models, are differentiated (a series of models instead of an integrated single model), the Narrative self is an integrated meta-model.

As the Narrative self is a product of the migration of selfhood, from the Social self through the self-model and then the Episodic self, it is more virtual than real; and yet it is the self we most often call on to define our me-ness. What is it that makes this self so attractive as a model of me? The answer appears to be that the Narrative self provides the individual with a sense of unity and purpose. It establishes the two cognitive concepts that having a self is supposed to enact: the concept of the single me and the concept of the continuous me. Although we cannot know it from the inside, this is what the Actual self seems to be from the outside: an entity delimited in both space and time; but, within those limits, a single integrated entity.

However, the Narrative self is also the most controversial of the selves. As we saw in Chapter 1, for Thomas Metzinger the Narrative self is an illusion because ‘we are not things, but processes’ (2003, 325). He is correct in saying this, inasmuch as our cognition is a process; but our brains and our bodies definitely are things. Our self can be seen as a system (a set of processes reliant on a particular structure to convert inputs into outputs). And a system contains both structure (a physical organisation which, when activated, converts inputs to outputs in a predictable way) and process (a particular route taken through a structure by an input to become an output). The Narrative self may be a virtuality, but it is also a product of realities that are emplaced in actualities. The Narrative self may not work as a thing; but as a metaphor or representation of a thing, it works just fine.

We have also seen that, for Galen Strawson, a Narrative self is not a prerequisite for being human. He says: ‘It is not true that there is only one way in which human beings experience their being in time. There are deeply non-Narrative people and there are good ways to live that are deeply non-Narrative’ (Strawson 2004, 13). Non-narrativity is, therefore, likely to be a correct diagnosis for some individuals without it affecting their cognition or socialisation. However, as Drummond et al. (2015) and Grossman et al. (2017) show, deficits in Narrative self do seem, at least in old age, to be associated with deficits in other cognition: the absence of a Narrative self in these cases is a by-product of other
cognitive conditions that cause narrativity issues more extensive than just the lack of a Narrative self.

When it comes to efforts to recreate a humanlike experience in machine form, scientists working in artificial intelligence have a clear understanding of the need for a Narrative self. For Pointeau and Dominey (2017), the Narrative self is a necessary tool for sharing plans, and it allows individuals to negotiate toward meaning in joint enterprises. These authors equipped their iCub robot with an AutoBiographical Memory (ABM), which is a simulation of a Narrative self, and showed that, when the ABM is linked to language, plans and activities could be negotiated between the robot and the trainer:

We previously suggested that shared planning could be developed based on 5 prerequisites: (1) object and agent perception, (2) perception of state changes (allows action perception), (3) ability to distinguish between self and other, (4) emotion/outcome perception, and (5) statistical sequence learning … These mechanisms, plus a specific ABM and methods for operating on the contents of the ABM allow for the capabilities reviewed in this report. As mentioned, we find the need for one additional capability, which is an interface between the language system and the ABM, in the form of a situation model. This is required in order to explicitly represent narrative relations between events that are not accounted for in the ABM.

(Pointeau and Dominey 2017, 16)

The six-feature ABM makes the iCub’s interaction with the trainer impressively humanlike. So we can see that, if Strawson is right that the Narrative self is not a necessity for a functioning human, or even if Metzinger is right that the Narrative self is not a thing, there is still a purpose for a Narrative self in the interpersonal negotiation toward meaning and joint enterprise; and a deficit in Narrative self makes that negotiation harder. It may be completely virtual, but there does seem to be some practical use in having a Narrative self.

The Cultural self: the self I should be

A Cultural self, like the Social selves, is a model offered to the individual by others; but, unlike the Social selves, it is a virtual self. It is a model of an ideal individual in this particular culture, explicitly the ideal self that the
individual can be. A culture usually has many ideal models, differentiated by gender, role, lineage, age group and any other way that the culture divides up its population. For instance, the Hindu caste system is based largely on gender and lineage, and it delimits not just the range of roles possible for an individual, it dictates how they are treated, whom they can marry, what they can eat and even what or whom they can touch (Pratheesh 2015). There are four main castes: priests, warriors, owning professions and labouring professions – a pattern repeated in internally specialist societies across the world. Unlike most other systems, however, the four Hindu castes are formally subdivided into sub-castes (in other cultures this level of differentiation is usually informal); but, like many other systems, there is also a formal gender-based differentiation, further limiting life choices. The caste system is a powerful engine for ensuring that life goes on regardless of who is in charge; but this also allows one ruling class to be replaced by another relatively seamlessly, without affecting the day-to-day functioning of the society. After seizing power in India, the Delhi Sultanate, the Mughal Empire and the British Raj all re-emphasised the caste system to retain control over the populace. The Hindu caste system is one of several historical systems that modern, global, pluralistic societies are breaking down; but a socially differentiated system nonetheless remains an important feature of most cultures today. The limited mobility between British social classes remains a case in point.

However, even if the range of ideal, or cultural, selves offered by a culture is quite wide, the options offered to each individual usually remain quite limited; and out of the small range offered, each individual often chooses one model as their lifelong Cultural self. This limitation is even observable in perhaps the least class-bound society in the world, that of the USA. Kraus and Park (2014) showed that perceived social class is correlated with self-evaluation: the higher the former, the higher the latter. Yet individuals do not usually change their Cultural self to enhance self-evaluation; the model they have been given is the model they accept. This acquiescence to the group opinion is even more noticeable in another modern culture, which places a high value on conformity to the Cultural self. As Markus and Kitayama (1991) say, ‘In America, “the squeaky wheel gets the grease.” In Japan, “the nail that stands out gets pounded down”’ (224). They describe the American approach as an independent view of the self – the cultural view of the self does not include other individuals; while the Japanese approach is an interdependent view of the self – the Cultural self is projected into the social calculus of the individual. The authors suggest that Western belief in the
Cultural self as independent seems to produce a less happy and healthy human society than the interdependent model.

Toon van Meijl (2008) pursues a similar idea when he describes the need for a ‘dialogical self’ to counter the limiting idea of a single Cultural self. Nowadays, we are faced not with a single culture to which we must relate, but a multiplicity of cultures we dip into, and aspects of which we incorporate into our Cultural self. The Cultural self is no longer a single target of selfness for which I should aim – the best me I can be – it is a changing and moving target. To keep up with the changes to the Cultural self, we need a dialogical self between the received social models and the received Cultural self. This seems to be missing from the SSMH model presented here, but this is because van Meijl has agglomerated the Cultural self with the Projected self. Where the self-model acts as a buffer between the received Social selves and the Projected self in the SSMH model, van Meijl places the dialogic self between the received Social selves and the Cultural/Projected self.

In the SSMH model, the main route by which the Social selves are incorporated into the Projected self comes around the other side of the ‘wheel’, through the self-model, the Episodic and the Narrative selves. The Cultural self (the self I should be) stands alone as a separate societal imposition on the self-model and on the Projected self. The best me I can be is a different imposition on my selfhood from the social calculus models of me: it is not how others see me but how others wish me to be. In the SSMH, there is no need to posit a novel mechanism to handle modern differentiated culture, because we could not have developed modern differentiated culture if the mechanisms for it were not already present in some form.

This does not mean, however, that modelling a Cultural self in a modern differentiated culture is simple or even linear. Navarro et al. (2014) showed that, for Mexican American college students, there was an iterative relationship between the Cultural self and the personal self (or self-model), such that retention of cultural heritage increased self-esteem, and higher self-esteem led to greater heritage-culture retention. Nataliya Aristova (2016) showed that changes in the Cultural self can also feed back out to the culture via the Projected self: the Cultural self projected onto the individual by the culture is, in the end, just the sum of the Cultural selves projected onto the culture by the individuals. As Aristova says, ‘Self-identification through culture and building up new cultural identities in new socio-cultural environments will always remain very significant factors for the self-determination of nations, countries and regions of the world’ (Aristova 2016, 160).
Chien-Ru Sun (2017) looked at a different issue regarding the Cultural self: the number of selves the culture imposes on the individual. Sun identified four types of self in the relationship between the individual and Chinese culture: the individual-oriented self (which collocates with the self-model), and three culturally imposed selves. The Cultural selves are: the models of my ideal self offered to me by individual others in one-to-one exchanges (the relationship-oriented self); the models of my ideal self offered to me by my family (the familistic [group]-oriented self); and the models of my ideal self implicit in my culture (the other-oriented self). Sun describes the other-oriented self as ‘the most undeniable’ (2017, 14): for an individual in a Chinese culture, the Cultural self plays a significant role in generating self-models, and the cultural-self-models of who I should be, as offered by the society around me, are more significant than for a Western individual.

The Cultural self, like the Narrative self, is a virtual self; it has only as much value as its human society is willing to give it. This can vary from very little (in societies valuing independent selfhood) to very much (in societies valuing interdependent selfhood). However, the Cultural self is also an aspirational self: each individual has an individual relationship with their Cultural self, which varies according to their need or wish to conform, whether they see their Cultural self as attainable and (at the subconscious level) whether the Cultural self is more or less fit for them than other self-models. It is in the Cultural self that dispassionate self-sacrifice begins, so the Cultural self is the key to a large number of our anxieties and self-doubts; the Cultural self supplies both the angel on our right shoulder and the devil on our left. As Aleksandr Solzhenitsyn put it, ‘The battleline between good and evil runs through the heart of every person.’

The Projected self: the self I want others to believe me to be

What is the purpose of accepting all the offered models of my self and generating yet more models of my self from them? If the only product of self-modelling is self-sacrifice, then it seems a poor return for a hefty cost in terms of cognition. Two reasons why self-modelling should benefit the individual have been explored so far: first, that a more complete social calculus allows better representation of relationships within my group; and second, that being able to share social models increases the range of cooperative possibilities with other members of my group. In both cases,
the fitness of the individual is enhanced by being part of a fitter group. Another reason why self-modelling is so valuable may lie in the fact that, as well as being able to model my self, I can project that model back into the world. I am no longer an it modelled by others to manipulate me; I am a she or a he or a they. This makes me a person with an agenda, with whom meanings can be negotiated. Negotiating with others toward meaning about others makes me an active player in communication – a me and a you; which means that the self-model I present to the world becomes part of the negotiation.

The Projected, or public, self is an emergent feature of the social models I receive, moderated through three routes: first, via self-modelling, Episodic self and Narrative self; second, directly via self-modelling; and third, via Cultural self. Or, to put it another way, the Projected self is an amalgam of my internal representations of myself and the expectations that others put upon me. However, there also seems to be a feedback loop outside of the self that allows the Projected self to affect the Social self: the self I want others to believe me to be can become one of the selves others believe me to be. Dianne Tice (1992) showed that, where a behaviour is performed as a public act (and, therefore, is available to be presented back to me as a Social self), it is more likely to moderate my Projected self than if it is performed as a private act. Where the individual has an ongoing interaction with a group of people, the more pronounced are the adjustments the received Social selves make to the Projected self.

Sedikides and Skowrons (2000) describe how the public self emerges from the symbolic self (an amalgam of the self-model with the Social, Episodic, Narrative and Cultural selves) to control the presentation of self to others. They see the public self as contributing to private self-knowledge through ‘reflected appraisal (i.e., seeing the self as important group members see the self)’ (Sedikides and Skowronski 2000, 100). However, relying on self-appraisal in this way is not an internal process: the only way I can be aware of how others see me is via the social self-models they offer me. For Sedikides and Skowrons, the symbolic self relies on the opinions we receive, and self-esteem (how much I personally value my self-model as a person) is dictated by the interpersonal evaluations offered by others.

Yet the Projected self is not just an unconscious product of received models of my self; there is conscious input to the model, too. While the relationship between the self-model and the Projected self is largely not under cognitive control – I automatically project a version of my self composed from the models of my self I receive from others – the input
from the Cultural self is aspirational. The Cultural self is a virtual model that is not me, but rather what *me* could be; and it opens the way for any virtual self-model to be used in the creation of a Projected self. For instance, Nystedt and Ljungberg (2002), in a series of experiments, found that the public self is on a continuum between style-consciousness (what I am presenting) and appearance-consciousness (how I am presenting); and the private self is on a continuum between self-reflectiveness (knowing I have a self) and internal-state-awareness (knowing I am not invariant). They showed that treating the public and private selves as continua fitted the data from their studies better than treating them as monoliths. The nature of the Projected self seems to be contextual rather than fixed. Nigel Rapport found that we also tend to see our Projected self as aspirational rather than fixed; and, in a biography of the painter Stanley Spencer, he described the artist’s Projected self as ‘an individual engaged in a life project’ (Rapport 2005, 60).

The manipulation of the Projected self by others has become an area of interest recently. For instance, Lambros Malafouris (2008) showed how the Projected self can be socialised via the Social self (self-model) in quite unusual ways, creating selfhood beyond the bounds of the Actual self. He describes how a signet ring from a Mycenean tomb encompasses both a personal concept of *me* within a possessive concept of *mine*, and a sociocultural concept of *him* within a possessive concept of *his*. The fact that this beautiful and valuable artefact was entombed with the corpse means it was treated as part of the individual after death: traditional ownership ceases at death, but tectonoetic ownership (the association of an object with an individual, and only that individual) does not.

Louis Rougier (2014) asked whether, in a modern marketing environment, advertising should be directed toward the consumer’s ‘real self’ (an amalgam of the Actual self and the self-model) or their Projected self. He took the view that both selves need to be addressed, but that ‘the product design, packaging and communication must first inspire the projected self’ (Rougier 2014, 4). By addressing the *should* factor in the received Cultural self (‘the best me I can be’), the aspirations of the ‘real self’ can be manipulated toward a more conforming Projected self, one that will buy the product because they *should* rather than if they *need* to. The Projected self in this scenario has become a subconscious promise by the Actual self to adjust the self-model to be more like the cultural self-model. This is only a re-emphasis of the group culture’s expectations about the individual, conformity to which is a natural result of the individual’s gaining fitness from living in the group culture; but it is nonetheless somehow disturbingly coercive.
Recently, Hazem et al. (2018) showed how the manipulation of the Projected self via the inputs of others to the Social self is a subconscious physical process: we create our own automatic self-censorship. When others notice us (by calling our name or by physical touching) it sensitises us to, and primes us for, other bodily inputs; and this priming increases our physical self-awareness and receptivity. Hazem et al. say that, in their experiments, ‘the effects of own name, social touch, and eye contact on bodily awareness were subtended by common brain mechanisms’ and that ‘knowing that the contact is directly created by another human agent is essential to the effect of social contact on bodily self-awareness’ (Hazem et al. 2018, 6). They see human self-awareness as being an iterative social process, with the group view of me affecting my own view of me, and my own view of me affecting the group view.

Looking back at Figure 8.1, a simple interpretation would be that the Projected self is the outcome of a process of cognitive self-definition, and provides the route by which I influence the group around me; but it turns out that the group influences me long before I can influence it. In evolutionary terms, this is all rather odd: I need to project my self onto my group to enhance my own agenda and limit those of others; and, for most species, if they do this in a subconscious, no-holds-barred way, using just their Actual self, then they have the best chance of achieving their aims. However, as a human, I am cursed – or advantaged – by sharing social calculus, which requires me to consciously model a different kind of self. This self is, in turn, not one self but many selves, all of which are either real or virtual and not actual. They are not me selves but they selves, where they happens to represent me; they are emic selves rather than etic selves, true because the individual and the group agree they are true, not because they are definitionally true; and they can be both differentiated and integrated, generating a continuously changing kaleidoscope of selves as my current Projected self. The self I project to my group as the real me is one out of many possibilities: a little bit of the me you tell me I am; a little of the me I believe I am; a little of the self I was; a little of the self I should, or hope to, be; and all heavily censoring the subconscious me I actually am. Yet, after all this, the Projected self is not the final product of the process; instead, it is part of a larger process where my projected model of me is further edited in the minds of other individuals in the group, and then offered back to me as a Social self – a self you tell me I am. This larger process is a continuous iteration through a life: it usually starts at about age 2, with a realisation that others are talking to me about me; and it usually only finishes with death.
... And there’s more: some other selves

So far, the terminology of selfhood has been shown to be quite diverse. We have seen a multiplicity of terms (the Actual self is also the individual-oriented self, the Narrative self is also a narrative identity, and the Projected self is also the public self). We have seen amalgamation of terms (the symbolic or private self includes everything except the Actual and Projected selves; the autobiographical self is the Episodic and Narrative selves combined; and the real self is an amalgam of the Actual self and the self-model). We have also seen subdivision of terms (the dialogic self recognises that we have more than one cultural-self-model, and those Cultural selves – at least, in terms of Chinese culture – include the relationship-oriented self, the familistic [group]-oriented self and the other-oriented self). And we have seen evidence that the SSMH offered here may not be the whole story (the tectonoetic self includes all the selves plus objects beyond the boundary of the Actual self). Yet there still remain other important subdivisions of selfhood that have not yet been explored here.

One of these is Ulric Neisser’s Five Kinds of Self-Knowledge (1988). The significance of his work in the literature on selfhood means that the SSMH needs to be reconciled with the self-knowledge model if it is to be taken seriously. The two models are not that dissimilar, although one important difference is that Neisser discusses knowledge and not models, and he treats his selves as aspects of a single selfhood rather than different ways of being a self. Neisser’s five types of self-knowledge map across the SSMH as follows:

• The ecological self is the self as a natural and physical object, a relationship with the actual world. This maps well to the Actual self.
• The interpersonal self is the self in communication with other selves, the self presented for others to communicate with. This maps to the Projected self.
• The extended self is the self that incorporates remembered and planned events involving the self. It maps to an amalgam of the Episodic and Narrative selves.
• The conceptual self ‘draws its meaning from a network of socially-based assumptions and theories about human nature in general and ourselves in particular’ (Neisser 1988, 35). This is a good description of how the Social self (a network of socially based assumptions) and Cultural self (theories about human nature)
work together, using the inputs of other people as fodder for our cognitively generated self-definitions.

- The private self is the self we create to explain our self to ourself, a conscious representation of our internal, and therefore exclusive, cognition. The first impression is that this maps to the self-model; but, for Neisser, the private self represents a more fundamental and real idea of selfness than the ad hoc and easily changed self-model. To represent Neisser’s view in the SSMH, the private self is better represented by an amalgam of everything except the Actual and Projected selves, so it includes the extended self and the conceptual self.

Thus, while there are some differences between the two models regarding what a self actually is, Neisser’s kinds of self-knowledge and the SSMH use similar tools to describe how humans generate their selfness. Other models of selfhood are less easy to reconcile to the SSMH. For instance, as we saw in Chapter 1, Baars et al. (2003) introduce the concept of an observing self: if there is conscious perception of sensory inputs, then there must be a mechanism inside the brain that intervenes in the cognition of sensing and inserts that awareness. That mechanism is the observing self. Vast amounts of our sensory inputs are ignored by conscious awareness: our body sends over 10 million bits of information to our brain for processing every second; and, of that, about 40 bits are consciously processed (Nørretranders 1991, 125–6). We are probably perfectly capable of operating without an observing self to generate conscious intervention; so the fact it needs to be posited to explain conscious awareness is significant.

The observing self poses problems for the SSMH. First, the selection of what is given attention must be a subconscious choice, but the giving of attention itself must be conscious; but how can something be both subliminal and intentional at the same time? Second, the difference between the Actual self and the other selves is that the Actual self is unmodelled. Modelling is a matter of attention; so where in the SSMH would an observing self be able to provide that needed attention? Third, how does the observing self interface between the Actual self and the other selves? These are questions that the SSMH perhaps should, but does not at present, address.

On the other hand, a different question may indicate that the observing self is not without its own problems: where in the act of observing is a self useful? Surely selfhood comes into the interpretation
of observation, not into the act of observing? If this is the case, then the observing self is an aspect of the unknowable Actual self, and thus not a self that can be modelled; it is beyond the ambit of the SSMH. Krauzlis et al. (2014) argue from neurological evidence that attention is not a cause of decision-making but an outcome; it is not a part of the interpretation of observation, just a conscious recognition of the decision made by a subliminal interpreting mechanism. Domenico Guarino (2018) argues that this provides neurological evidence for Daniel Dennett’s (2009) ‘strange inversion of reasoning’, thus linking together cognitive attention as an evolutionary outcome with most other evolutionary outcomes: attention and evolution occur because an event produces an outcome that becomes significant, not because the event itself is significant.

Another approach is the patterned self, proposed by Shaun Gallagher (2013a). This model of the self is more deeply differentiated than the SSMH, but the SSMH largely corresponds with the patterned self in terms of form. What differs is that, where the patterned self is a single, coherent self that can present different patterns generated from a set of innate and personal psychological aspects, the SSMH treats the Actual self as singular but subliminal (and therefore consciously unknowable), and the consciously modelled selves are disparate products of received models of the self. The aspects of the patterned self include: minimal embodied and minimal experiential aspects; affective and psychological aspects; intersubjective, narrative and extended aspects; and situated aspects. However, within these aspects, it is not clear how much is under conscious control and how much is subliminal; and this matters, because humans are both slaves to our genes and controllers of our cognitive destiny. Any theory of selfhood has to address the fact that we have both selfness and awareness of selfness (Edwardes 2014).

Of the two models, the patterned self seems to be intuitively closer to how we believe our selves to work: we think of our self as unitary and integrated, with every social projection of our self being in some way honest and faithful to that integrated self. However, we also recognise that our single, integrated self is mutable, so the projection we make today need not be the same as the projection we made yesterday; but we also intuitively believe that today’s and yesterday’s projections remain faithful representations of the integrated self. In addition, we recognise that we have the capacity to deceive (both others and our own self) when we socially project our self. This is a lot of things to simultaneously expect from our unitary and integrated self, which indicates that treating the self as unitary may be intuitive, but it may not match the way selves actually work.
Another self is the culturally evolved self, proposed by Lloyd Hawkeye Robertson (2017). He takes the view that free will is an emergent outcome of being immersed in a complex cultural environment, and self-awareness is an emergent feature of free will. He does not address the question of how a complex cultural environment could emerge or evolve, or why it should emerge or evolve in the particular case of human socialisation; and he does not convincingly explore why complex cultural environments should be exclusive to humans, although he assumes them to be so. There is also a hint of circularity in his argument: our culture gave us free will, and our free will gave us self-awareness. However, he also says that ‘Free will originates from the first person experience of the world and one’s self-constructed understanding of his or her agency in the course of action’ (Robertson 2017, 4). So, if self-awareness comes from free will, and free will comes from first-person experience, what is the difference between self-awareness and first-person experience? And, if there is no difference, where does free will come in? Robertson raises important questions about the relationship between free will, culture and selfhood, but he may not have them correctly sequenced to provide answers.

Robertson also says that, ‘Questions of free will could not have been asked by beings unable to visualize the concept’ (2017, 6). This reminds me of the exchange between Alice and the Red Queen:

Alice laughed. ‘There’s no use trying,’ she said: ‘one can’t believe impossible things.’

‘I daresay you haven’t had much practice,’ said the Queen. ‘When I was your age, I always did it for half-an-hour a day. Why, sometimes I’ve believed as many as six impossible things before breakfast.’

(Lewis Carroll 1872, Chapter 5: ‘Wool and Water’)

Indeed, the capacity to believe impossible things would seem a much more reliable definition of humanity than either free will or self-awareness. We live in a world of solutions that were, at one time, impossible (metal ships, aircraft, cybernetic limbs, nanobots – the list is long and growing); but the problems were worried-at and debated until solutions were finally visualised and shared with others.

Another type of self that is not properly explored in the SSMH is what we can call the e-self. This is a version of my Projected self that is not subject to the knowledge others already have about me; it can therefore be more exploratory and less constrained than other Projected selves. The
e-self is not really a new phenomenon, and it has been described as an authorial self or authorial persona when discussed in relation to printed publications (Hyland 2001). However, the rise of electronic media and the appearance of the ‘casual’ (rather than ‘packaged’) author has meant that many more people are producing Projected selves directed at people who do not know the projecting individual except through the media. The Projected self in these cases is unverifiable, so one downside of the increasing technical sophistication that has made electronic media possible is the appearance of blocks of code complex enough to imitate a human e-self. They are often treated as actual humans, and they can be used to manipulate the gullible and cause havoc in a modern electronic democracy (Deb et al. 2017). This aspect of e-selves will not be further explored here, but the significance for the study of human selfhood is extensive.

The e-self is mostly unverifiable and uncensored. This means the individual has the potential to be brutally honest in their projected e-self, putting forward an image that they would never dream of presenting in face-to-face contact with longer-term acquaintances. Hu et al. (2017) show that, in all forms of communication where the individual’s identity is known and their reputation is at stake, the individual projects a positive self, which agrees with their positive self-model (the ought self) and their Cultural self (the ideal self). By contrast, in cases where the individual believes they are anonymous, they project more of their socially negative self-model and mostly ignore their Cultural self. For some people, this can have the effect of making communication in the anonymised online environment more attractive than in the actual reputation-driven world: some individuals seem to derive greater satisfaction from online anonymous communication than from any other form of human contact.

On the other side, Gil-Or et al. (2015) showed that e-communication makes it easier to project a deceptive self-image. This deception involves not just the daily white lies of normal, reputation-driven communication; it allows the individual to alter major features of their Projected self – age, gender, beliefs – anything, really. This creates what Gil-Or et al. call a false Facebook self, which can become a pathology in itself, or can license other psychopathologies the individual may have. Because of this, the creation of a false Facebook self is not just a personal choice, it can be a social problem; and the authors recommend that steps be taken to identify and suppress these pathological false Facebook selves. We are only now beginning to address the serious social issues that the projected e-self creates.
The final approach to human selfhood examined here was provided by George Lakoff (1992) with his multiple selves. Lakoff approached the question in a very similar way to this book, giving a series of conceptual metaphors to show how human selfhood works; and I must acknowledge Lakoff’s inspiration for many of the ideas I have presented. He proposed a binary approach, with the Subject consisting of consciousness, will and judgement, and the Self being the rest of the person. However, there is no single way in which the Subject interfaces with the Self, creating different types of self-models within the individual’s Subject+Self cognition. The first of these is the Projectible Self Model, which is not the same as the Projected self; instead it is the capacity to produce a Projected self, using tools like pronominalisation. The second is the Objective Subject Model, which allows the Self to perceive the Subject as an object – somewhat like a self-model, but once again more about the capacity to produce a self-model than the self-model itself.

The Objective Subject Model gives Lakoff his first conceptual metaphor: KNOWING IS VIEWING allows us to treat our subject-model as if it were a thing, because we are consciously aware of ourselves as subjects. Two other conceptual metaphors Lakoff presents in relation to the Objective-Subject Model – ENHANCED CONSCIOUSNESS IS THE ABILITY OF THE SUBJECT TO SEE THE SELF FROM THE OUTSIDE and OBJECTIVITY IS THE SEPARATION OF THE SUBJECT FROM THE SUBJECT’S VALUES AND PRESUPPOSITIONS – are more definitions than metaphors, so will not be considered here.

The third way the Subject interfaces with the Self is the Separable Subject Model, which reverses the Objective Subject Model, allowing the Subject to perceive the Self as an object. Lakoff sees this as generating a series of metaphors, such as LACK OF NORMAL CONSCIOUSNESS AND CONTROL IS BEING OUTSIDE THE SELF, RETURNING TO NORMAL CONSCIOUSNESS IS COMING BACK and EUPHORIC STATES ARE UP. This leads on to the Scattered Self Model, which reflects the fact that the Subject can maintain several different selves (or several different loci of self) at the same time – for instance ‘I’m all over the place today, I don’t know whether I’m coming or going’; and the Scattered Self Model leads to Loss of Self models, where the self is separated from the Subject and even from other selves.

Lakoff and Johnson (1999, Chapter 13) revised the model, proposing six selves: the Physical-Object Self, a self we can manipulate; the Locational Self, the self situated as an object in space; the Scattered Self, the result of trying to maintain several different self-models simultaneously; the Social Self, the self as a product of, and link to, other selves;
the Projecting Self, the self we present to other selves; and the Essential Self, a conscious representation of the unconscious Actual self. Lakoff and Johnson’s six selves have some correspondences with the SSMH, but there is one important difference: Lakoff’s Multiple Selves are all conscious selves, whereas the SSMH includes the subliminal Actual self. This difference gives the SSMH, an origins model, a physical base on which the other aspects of selfhood can be anchored and from which they can be built; the Multiple Selves model, an embodied model, does not need this.

**Why self defines us**

What defines a self? Thagard and Wood (2015) produced a list of 80 self-phenomena, grouping them into six main classes: self-representing (oneself to oneself, oneself to others, and evaluation); self-effecting (facilitating and limiting); and self-changing. This list of phenomena shows that, whatever a self is, it is a widespread feature of human cognition, being involved in sub-phenomena such as compassion, forgiveness, reliance, effacement, deception and realisation. The widespread effects of selfhood in cognition mean that a simple, one-line definition of self is probably impossible, and very likely deceptive.

Significantly, Thagard and Wood place consciousness under the experience of self-representing oneself to oneself: in other words, they see awareness as a cause or outcome of selfhood. In this book, we have separated other-awareness and self-awareness, meaning that awareness can be treated as both a cause (other awareness) and an outcome (self-awareness) of shared selfhood: selfness and awareness of selfness emerge from the sharing of awareness of others through language. Thagard and Wood show that having a self, with its concomitant self-phenomena, is a defining feature of humans – both in terms of its source (social calculus shared through language) and in terms of its outcome (the enhanced cooperation of joint enterprise).

A self is a useful thing to have: it allows us to show the persona we wish to show to the people around us. However, what makes it really useful is if other people have a self, too: when humans relate together through their Projected selves or personae, rather than their personalities, they create a pragmatic layer in their communication. This pragmatic layer acts as a lubricant between personalities, allowing them to cooperate efficiently enough to engage in complex joint enterprises. Human language is full of expressions that represent this pragmatic layer at work; for instance, ‘don’t take this the wrong way, but …’ (a way of
criticising a persona without involving the personality) or ‘let’s do this’ (a depersonalised way of generating provisional agreement between personae) or ‘so what do we do now?’ (an invitation to the other persona to offer a solution for consideration, not action). Our Projected selves are able to work together in ways that our personalities cannot.

Personality is not a human-only characteristic; many other animals display identifiable differences between individuals. Daniel Nettle (2006) reviewed several very different species, including birds, fish and insects, and found evidence of personality in all of them. Seyfarth et al. (2014) show that, while kinship is a major feature in determining rank for baboons, having a less aggressive, more conciliatory personality also has an effect on the individual’s rank. Baan et al. (2014) found the same was true for wolves. However, to date, there has been no evidence for persona in any other species. To select a persona to project, an individual needs to be aware of themself as a self – which we humans can do only because we share our social calculus using language. So language, shared social calculus and modelling a Projected self all seem to be markers of being human. Having selves is what makes human culture what it is: the reality of each individual self, both for others and for the individual, makes our self-driven society (as compared to the ego-driven society of chimpanzees) an actuality.

However, while the existence of a human self-driven society is an actuality, that society is composed of self-aware selves. Self-awareness is a modelling process; so, while the process of modelling can be an actual thing, the models themselves are mental constructs, which are virtual. We can agree to treat them as if they were real, but we cannot give them actuality with just our belief. This is where the link between the personality (an aspect of the Actual self) and the persona (the Projected self) comes in: inasmuch as the Projected self is treated as a personality by others, it is a proxy, or a metaphor, for an actual thing. This still leaves a dilemma: is the reality of selfness enough to prove that the aware self is not an illusion or delusion? The answer relies on how we define selfness, and what we want selfness to say about us. If selfness is a way of explaining how humans are able to work together, then the fact that we engage in actual joint enterprises is enough to justify selfness as a necessary part of that process. If, however, I am looking for my self as a concrete example of selfness, then I am on a fool’s errand.

One feature of the SSMH still needs to be addressed: the hypothesis relies on an ability to take models of my self offered by others and treat them as if they were third-person models of me. This means we would expect that modelling others and modelling me would be similar
cognitive experiences, using many of the same cognitive processes. Fortunately, this is just what Francesca Happé (2003) has found. She looked at nine brain-imaging studies conducted by seven teams to assess brain usage in ToM tasks, and found that the same areas were active in both self- and other-modelling. In her conclusion she says, ‘neuroimaging findings to date appear to suggest a network of regions involved in attribution of mental states to others which largely overlaps with areas of activity in self-reflection tasks’ (Happé 2003, 141). It seems that, as the SSMH predicts, modelling others and modelling me are treated by the brain as similar processes.

In summary, therefore, we can say that having selfness is being human; and that is why self defines each of us – both from the outside, with others’ models of me, and from the inside, with my models of me.

Notes

1. There and Back again is the second half of the title The Hobbit, or There and Back again by J.R.R. Tolkien – which starts with a map of the journey. There and Back again was missed off the cover of the first edition (George Allen & Unwin, UK, 1937), but has been reinstated in some subsequent printings.
2. The Standard Statistical Brain (SSB) is a convenient model of language in the brain, which corresponds quite well to about 95% of actual brains; the other 5% can differ quite markedly from the SSB without the language capacity being affected. The SSB is a useful tool; problems only arise when it is seen as a real standard brain, an actual standard brain or, worst of all, an ideal standard brain.