Building up to selfhood

An anthropological perspective on where selfhood came from, and the selves humans have

Martin P.J. Edwardes

A presentation in the series *Breaking Down the Self* for The Open Foundation
Thursday 17 June 2021, 19:00-20:00
Also for the presentation *Being Someone: Language and Selfhood*
Friday 30 April 2021, 13:30-14:00

Contents

Introduction

When asked, almost every human being will tell you that they are an individual. They are a unique version of being human because nobody else has had the experiences they have had from the same perspective as them. They have, inside their heads, an exclusive approach to being, because they are the only person able to be them. We say that they have a self … no, they *are* a self … and being that self excludes them from being any other self; and it also excludes another self from being them.

But what is this strange substance of selfness, and why is it simultaneously a commonly shared commodity and a totally personal feature of the individual? What does it mean when I say I have a self, or am a self; and does having a self inevitably also mean that I know I have a self? Our personal relationships with our selfs would seem to indicate that most humans do, indeed, know we each have a self: conventional wisdom is full of nostrums about self-knowledge. For instance, over the entrance to the temple of the Delphic Oracle was inscribed, "know thyself"; Shakespeare said, "to thine own self be true", in the belief that we would understand what he intended; Descartes said, "I think therefore I am", in the belief that this defined his selfhood; and Jerry Herman said, "I am what I am", and set it to music. It seems that knowing your self is a recursive, or at least iterative, process wherein the self can know the self knowing the self.

This, however, is a strange thing to be able to do; knowing my self means treating my self as a thing external to my self – a third-person thing rather than a first-person self. If a fellow human sensibly follows Darwinian strategies (whereby the genetic self is treated as more valuable than any other genetic being, at least as a vehicle for getting the self's combination of genes into the future), then we collectively denigrate them as selfish; and if they do the evolutionarily insane thing and put the welfare of others above their own, we collectively praise them as selfless. Yet, as individuals, we can only act selflessly if we are able to compare other selves with our own self, and we can only do that if we have a conscious knowledge of our self – or, at least, a conscious knowledge of a cognitive model of our self. Having self-knowledge seems to be a socially useful and desirable thing; but, if it makes us less genetically fit, is it actually an evolutionary aberration?

This presentation looks at the origins and nature of human selfness to attempt to answer some of these questions. The content of this presentation is taken largely from the book, *The Origins of Self: An anthropological perspective* (Edwardes, 2019).

The peculiarity of human selfhood

Before looking more closely at selfhood, we need to consider two related features of being human: our ability to model the interpersonal relationships we have with others; and our ability to share our models with those others. Chimpanzees share with us the ability to maintain a social calculus of the relationships between others in their social group (Call et al, 2004): they are able to anticipate the emotional states of other chimpanzees – and, indeed of other humans – and adjust their behaviour appropriately. However, unlike us, they do not seem to exchange that social calculus with others (Whitehead, 2001). In contrast, the human communication system, language, seems to be

ideally suited for this type of exchange, and that may even be the reason why we have it (Dunbar, 1996).

The social calculus a chimpanzee needs to navigate their group is a two-argument form that we would recognise as a human sentence: [person A] [has a relationship with] [person B]. This can be expressed more simply as a linkage between two nodal models of other individuals, or node-link-node; and it can also be expressed as a grammatical noun-verb-noun, or subject-verb-object relationship. The cognitive structures needed to formulate these two-argument forms can be described as differentiation and segmentation: differentiation allows us to see the two arguments or nodes or nouns as being of a different nature to the relationship or link or verb; and segmentation allows us to identify the two-argument form as a proforma, so that any argument or node or noun can occupy each of the two argument slots, and any relationship or link or verb can occupy the relationship slot. Chimpanzees have the cognitive structures needed for social calculus, but they do not have the need to communicate them; we do, which could be the reason why our communication system closely maps to our social calculus cognitive structures.

Humans can and do share relationship models, and we can interpret the models shared with us. In this way we can supplement our own social calculus with any reported relationships we receive. We have to remember, however, that a relationship model offered to us by another individual does not have the same valency as a relationship model generated by personal observation: the offered model may be less current than my own model, or it may reflect other features of the relationship between the offeror and the people modelled, or it may even be deliberately deceptive. A capacity to tag received models with the identity of the offeror therefore becomes useful to the receiver; and to do this a three-argument form becomes necessary: [person C] [says that] [[person A] [has a relationship with] [person B]]. This three-argument form is hierarchical as well as differentiated and segmented: the two-argument form is encapsulated within an attribution, which gives us the third argument. It also means that an individual's social calculus can integrate less reliable received relationships with more reliable personally observed relationships, because their truth-values are established by their hierarchy.

Once we have three-argument models in our calculus, it becomes possible to share them by extending the grammar of our language to include the extra grammar of our social calculus. We could simply share received models as two-argument forms, but that could leave us open to accusations of inaccuracy or falsehood: if I tell Ethan that "Beth likes Colin" because Alf has told me that Beth likes Colin, then I risk my reputation with Ethan if Alf is wrong or lying. Fortunately, having the three-argument form cognitively available, I can explicate this instead of the two-argument form: "Alf says Beth likes Colin"; and the recipient of this three-argument form, if they have been exposed to offered two-argument forms, should have three-argument grammar available to deal with this. Ethan should now understand that the relationship information I am offering is not guaranteed by my reputation but by Alf's, and can treat it accordingly. Additionally, because hierarchy allows "thinking in levels", Ethan can cognitively map the relationship information as

"Doris says that Alf says that Beth likes Colin", and then pass it on. This allows for the fourth cognitive structure of language: recursion. Theoretically there is no limit to the number of [X says that] nestings possible, although practically it seems to break down after five or six iterations (Dunbar, 2004, ch3).

By passing on a received model hierarchically, the speaker need take no responsibility for its accuracy. If the new receiver wishes to use the information then they have to pay the cognitive costs of attribution. This means that the exchange of social calculus models becomes more than a dichotomous passing of true or false utterances; instead, every utterance needs to be treated as the Russian proverb advises: trust but verify. Language is not about facts, it is about opinions, and this is what makes it such a good vehicle for fiction: it is the lie you tell me that we both know is a lie, but which I nonetheless treat as if it is, on some level, true (Boyd, 2017).

The significance for selfhood of the sharing of social calculus can be shown in the answer to the question, what happens when the receiver of a social model is also a protagonist in the model? When Alf offers me the social model "Beth likes Colin" and I am Colin, then I have a dilemma. Prior to the offering of this model in which I am a protagonist, I would have had no use for a conscious model of myself: self-knowledge is a social construct, and only needed if I am to model my self for comparison to other selves (Brown, 1997). However, if I am incomparably different from other selves – because I am unconsciously present in my cognition as the one and only first person, and not consciously modelled as a third-person entity – what is the value in modelling my self? The presentation by Alf of his model of me means that I now need a new node in my social calculus, a node which is no different from the other nodes (a third-person entity), but which is also a conscious marker of my own existence. To model myself as a third-person entity, I must be aware of my self as an entity.

This side-effect of sharing of social calculus information is discussed in more detail in chapter 5 of *The Origins of Self*, but for the purposes of this presentation we only need to understand that humans are able to share social models because we have the communication system and the willingness to do so.

The development of selfhood

Let us turn to the ideas in chapter 2 of *The Origins of Self*. This chapter looks at how life has dealt with selfhood from single-celled animals to modern-day humans. This is an ambitious timescale, but necessary to understand human selfness in context. Fortunately, there are only five key stages in the development of our species-specific selfhood, and only eight strategies of self to be considered:

1. The evolution of a **Sense of not-self** (strategy 1)
2. The evolution of a **Sense of almost-self** (strategy 2)
3. The evolution of a **Sense of other** and a **Sense of self** (strategies 3 and 4)
4. The evolution of an **Awareness of other** (strategy 5)

5. The evolution of **Joint enterprise**, and the emergence of **Awareness of self** and **Awareness of selfness** (strategies 6, 7 and 8)

The eight strategies will each be considered in more detail in relation to the five developmental stages (all are further described in the glossary of *The Origins of Self*).
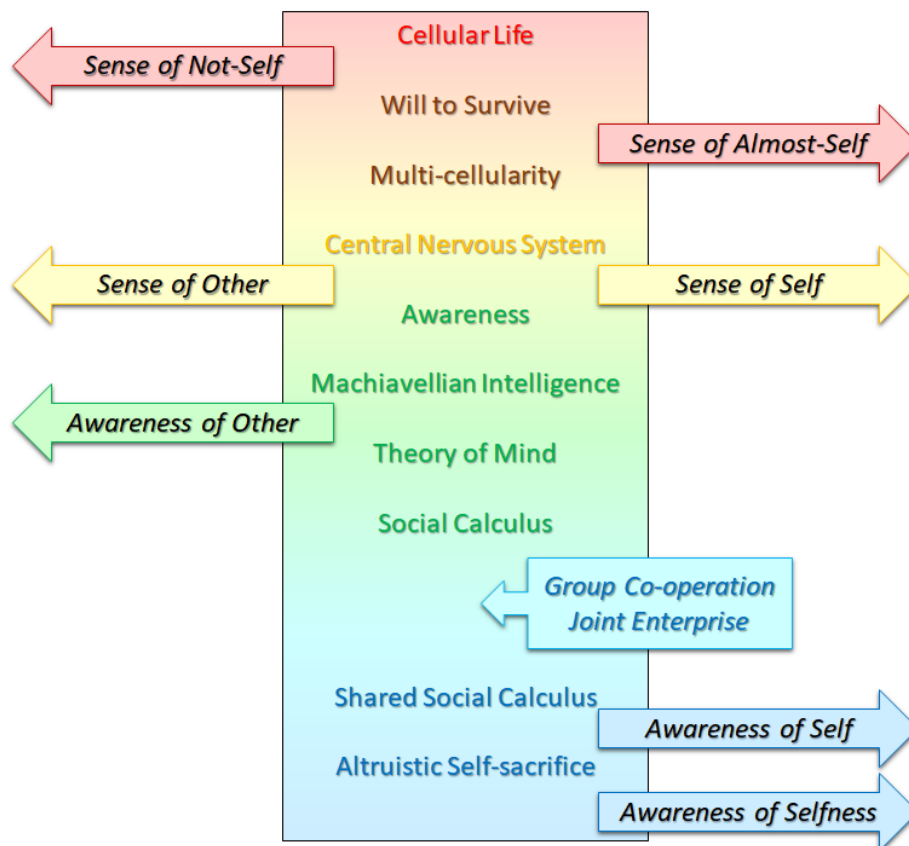

Figure 1: The development of selfhood

Sense of not-self

The first strategy of self that is likely to have evolved is a **Sense of not-self**, which seems to have occurred early in evolutionary terms: we can see it at work in any single celled-animal, such as an amoeba. Functionally, an amoeba is a set of biological mechanisms within a flexible cell membrane, and this membrane marks the boundary between the internal ends of existence (movement, feeding and division) and the external means by which that existence is maintained (identifying predators to be avoided and prey to be consumed). Things outside the amoeba's cell membrane have the potential to support or damage its existence; everything inside it is (theoretically) already devoted to the amoeba's existence. (This is not always the case, viruses and other inimical intruders can enter the cell and hijack its existence; but, once inside the membrane, the invader is relatively safe from any defensive measures by the cell, and the cell is usually doomed (Padariya et al, 2021).

Because of this outside-inside difference, it is useful to the amoeba to be able to detect items which are outside its cell membrane and react appropriately to them; and this gives it, for all practical

amoeboid purposes, a Sense of not-self (Burton et al, 2021). However, this is not on any level a sense of which the amoeba needs to be aware.

Sense of almost-self

When clonal colonies of cells begin acting as if they are multicellular organisms (as slime moulds seem to do), then a **Sense of almost-self** is needed (Boussard et al, 2021). This is a recognition that the cells around a cell may well be clones of the cell, so treating them as not-selves is literally self-defeating. Cells which are members of the clonal colony need to be treated differently from non-members, creating an "us" and "them" relationship with the universe. This extra sense also gives a cell another route to get its genes into the future: assisting clones to survive has exactly the same fitness value as individual survival, because the clone cell has the same genetic make-up as the original cell (indeed, as the cells replicate by genetic division, both daughter-cells can be viewed as the original cell). The individual cell is no longer the primary vehicle for reproduction, it is the colony that matters. Sense of almost-self therefore leads toward multicellularity, where individual cells can specialise to better support the colony. A multicellular clonal colony can behave as if it were a single entity, and not just composed of entities, which creates a new way of being a self. However, this new self remains unconscious of its selfhood; awareness is still not needed to have a Sense of almost-self.

Sense of other and Sense of self

When clonal colonies begin co-operating then cellular specialisation can begin. While all cells retain the possibility of adopting a range of strategies, it is not necessary for all cells to adopt the same strategic response to an environmental cue. For instance, most slime moulds exist as individual cells when food is abundant; but when food becomes scarce they aggregate to move to new hunting grounds. This process can often lead to sporulation, whereby some of the individual cells become spores, which can be picked up by the wind and dropped into new environments, while other cells clump into stalks, which lift the spores up into the wind. The spore cells have a chance at continued survival, the stalk cells do not; but, as they are clones, the survival of the spores equates to the survival of the stalks.

Over a period of about 2 billion years, this functional specialisation of individual cells led to multicellularity. With multicellularity, the vast majority of individual cells in a clonal colony adopt roles which are so specialised that they cannot survive individually. Muscles, bones, nervous systems, and a host of other specialised systems develop into what we treat holistically as a single animal; and, like an individual cell, an individual animal is a system for receiving inputs from the universe and converting them into internal resources. It is the means and ends distinction at work again, but at a higher level of organisation.

From multicellularity come two new strategies of self: **Sense of other** and **Sense of self**. Sense of other allows an animal to treat other animals as entities with agendas, thus differentiating the universe into (living) things with agendas and (non-living) things without agendas; and Sense of self

allows the colony of cells composing the individual to behave as a single individual with a singular agenda. Sense of other and Sense of self map to the cellular level senses, Sense of not-self and Sense of almost-self. However, they have different roles: where Sense of not-self is a recognition of the cell wall as an existential boundary, Sense of other is a recognition of things beyond the boundary; and where Sense of almost-self is a recognition of the transactional nature of the cell boundary, Sense of self is recognition of a new existential boundary between the individual animal and the rest of the universe. Neither Sense of other nor Sense of self require an Awareness of other or self, they only allow the individual animal to subconsciously recognise that there is a difference. However, they do both seem to require, or at least work better with, a central nervous system.

Awareness of other

The next developmental event is that the clonal colony, or animal as we can now call it, develops awareness of events outside the individual animal – **Awareness of other**. The animal becomes capable of consciously adjusting its activity to accommodate, or deal with, those events. The animal develops awareness that things with agendas can and do react to the animal's activity, which makes them manipulable. From our Homo sapiens viewpoint we would say that things with agendas can now be consciously treated by the manipulator as other individuals, but that is not a capacity the manipulator actually needs at this stage. The manipulation of others we describe as **Machiavellian Intelligence**, precisely because it treats other entities as manipulable things with agendas and not as intentional individuals. However, the manipulation of others is more successful if the manipulator can also deploy a **Theory of Mind**, treating the other animals as cognitively reactive (having a mind), and thereby modelling those others more predictably. Theory of Mind, in turn, leads to social calculus, the capacity to model the effects of different manipulations in terms of advantages and disadvantages to the self. By this stage, the manipulator has a sophisticated awareness of others as intentional beings, but it still has no need to be aware of itself as intentional: it only needs to be aware of others as intentionally reactive, it does not need to be aware of its own awareness.

Awareness of self and Awareness of selfness

From Theory of Mind and social calculus, the human clade took an unusual turn into a sixth strategy of self: **joint enterprise**, **group co-operation**, and the peculiar communication system we call **language**. Whether language originated as a strategy for sharing social calculus models or for some other purpose, it is so effectively suited for describing social calculus models that its early adoption to share them is very likely. This, as we have seen, led to some interesting cognitive conundrums concerning truth and selfhood; and these, in turn, led to the breakdown of the new existential boundary given by Sense of self: the boundary between the individual and other individuals. With joint enterprise we began behaving as components in the larger system of society – just as, earlier in evolution, clonal cells began behaving as components in a larger multicellular system. Mark Twain famously said, "history doesn't repeat itself, but it rhymes"; it seems this applies to evolutionary history, too.

Two evolutionary strategies developed from the revolution of language, the first being **Awareness of self**. This is the realisation that if others are offering me models of me then I can also model me. To do so, I need to treat my newly-acquired model of me in the same way I treat my models of others: as a third-person entity and not as a first-person event. This means I have awareness of my model of me in the same way as I have awareness of my models of others. This is not self-awareness in the classic, Delphic view ("know thyself") but it is sufficient awareness of self to give me the illusion that I know myself.

Thus we come to the final strategy of self: **Awareness of selfness**. Being able to make models of myself makes me aware of the mutable nature of my Awareness of self: it is based not on a single physical entity but on the many models of myself that I can build, which are, in turn, based on the many models of myself I receive from others. This Awareness of selfness is still not Delphic self-awareness; but it makes it possible to believe that, somewhere behind all those models, an actual self exists. It is this Awareness of selfness that seems to create a lot of the problems of selfhood that can afflict humans; but that is a tale for another time.

The Seven Selves Hypothesis

With Awareness of selfness we have achieved the full "suite" of Homo sapiens social modelling capacities. However, the way we use those social modelling capacities to make models of our self, and to represent our self both to ourself and to others, is another matter. Modelling is a constantly ongoing iterative process which has no final product, only snapshots of the current state of the system; and the part of the system you inspect for a snapshot affects the representation you get. *The Origins of Self* presents, in chapter 8, a hypothesis where seven different types of modelled self are possible at any time; and in some cases, multiple copies of the same type of self are also possible. This description of selfhood is rather unimaginatively labelled the **Seven Selves Hypothesis** (or SSH).
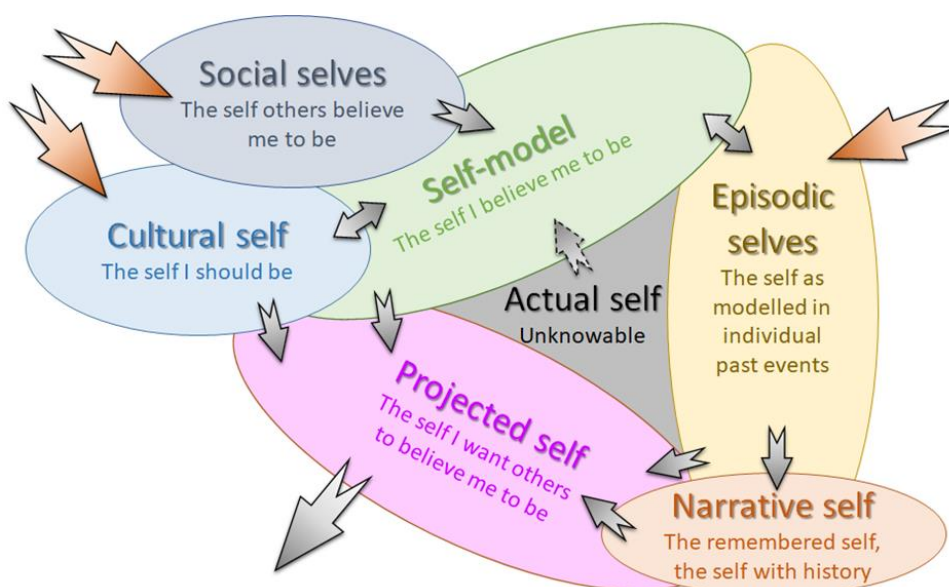

Figure 2: Seven Selves Hypothesis

Types of self

In figure 2 we see the seven types of self as proposed by the SSH, two of which are expressed plurally (Social Selves and Episodic Selves) while the others are expressed singularly. This is not intended to indicate that we each have only one Self-model, one Narrative Self, one Cultural Self and one Projected Self; instead, it is to indicate that we each can be aware of only one of these selves as being "me" at any one time. I can hold a Self-model of who I believe I am and simultaneously compare it to Self-models of who I can be, and I can swap between them over time; but I cannot be two of them at once. In contrast, I can hold multiple models of my episodic selves without any of them being the current "me" – because they are all models of who I was, not who I am; and I can hold multiple models of my Social Selves because they are all models of other people's opinions. The exception to this modelling of selves is the Actual Self: its nature is different from those of the other selves.

The Actual Self: Unknowable

The first self we have is the **Actual Self**, which is described in figure 2 as unknowable. This is the physical self without which there can be no selfhood, and it precedes human awareness of self by a large margin: it was there with the first multicellular animals and equates to the Sense of self (see figure 1, The development of self). The Actual Self is a product of the Darwinian genetic maxim of survival through fitness; and, unlike the other selves, it is not a conscious representation of the self. Freud would associate it with the id, of which he said,

> "All that is true is that everything that happens in the id is and remains unconscious, and that processes in the ego, and they alone, can become conscious." (Freud, 1959 [1926], p197).

Yet unknowable does not mean unmodellable: it is certainly possible for me to consciously model what I believe is going on in my subconscious. However, the model being built here is just another Self-model, so it is impossible to generate a model of my Actual Self before I have the capacity to Self-model; and what is being built has no conscious input from the Actual Self, so it is therefore unlikely to be a usefully accurate representation of my Actual Self. This means that the ability to model a semblance of my Actual Self is just a side-effect of being able to Self-model: it is essentially illusory and likely to be self-deceptive.

The Social Selves: The self others believe me to be

The first type of self of which I am consciously aware is the **Social Self**. This is a model of my self which has been offered by others as part of the sharing of social calculus. Unlike most of the other selves, it is provided wholly from outside the self. My Social Self selfhood is not modelled on my internal representations, it is, to use Gould & Lewontin's (1997 [1979]) term, a spandrel of socialisation. Nonetheless, the Social Self is the start of human self-awareness, and the first "me" that "I" am aware of.

The Self-model: The self I believe me to be

From the Social Selves that I receive in shared social calculus, I am able to build up a reliable picture of how others see me; I can begin to form an integrated model of who I am. This integrated model, however, remains an amalgam of differentiated and often disparate other-models from other people, so it is integrated but not unitary; and, because the expectations of who I am change as the people around me change, my currently active **Self-model** changes to match the social context; which, in turn, means that the self I perceive as me is constantly changing 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? This fraying of selfhood can be seen in several psychological pathologies, but it is not a topic pursued in this presentation.

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

The next selves we become able to model are the self-as-was and the self-as-will-be. Traditionally these are referred to as **Episodic Selves**, because they are selves remembered from the past and selves modelled into the future; they are episodes in a continuing and unitary life, generated by the activation of self-memory within the self-modelling mechanism. The activation of memory, however, introduces the standard issues involved in remembering. We tend to think of remembering as the faithful retrieval of engrams from a store of fixed memories, a store which may degrade over time but consists of, in the short term, mostly reliable recordings of events. We now know that this is not how memory works (Gravitz, 2019).

When I remember an event, I rebuild the event from fragments that come to conscious attention. Most of these fragments are either directly from the event (triggered by the cognitive "X" that began the remembering process) or relevant to the event (secondarily triggered by a remembered fragment); but some are intrusions – not related to the event. Most of these intrusions are consciously pruned out, but some get added to the memory. Also, some aspects of the memory are not recalled because they are not triggered. Even in the short term, my memory of the event is selective and incomplete. In addition, my episodic selves are not mine alone: other people have their memories of events in which I was involved, and their memories, when shared, can influence my memories.

This unconscious editing becomes worse when I am trying to place myself in the memory – even if just as the observer. When I remember my self, I do not remember my Self-model as it was when the memory was laid down; rather, I construct a current Self-model to represent my previous self. To use a metaphor from electronics, recalling a memory involves copying the memory-engram into working memory, adjusting the memory to make it a memory of my current Self-model, and then writing it back – usually by overwriting the previous engram. Every time we remember, what we remember is an edited version of the memory; and the more frequently we recall, the more edits creep in.

Between these several effects it seems that we cannot stop ourselves from continually counterfeiting our memories. However, while this is less than ideal for a legal system which relies on accurate recall and faithful retelling, it is useful for the rest of us. Memory of self is not inherently fallible just by accident; it is important that it be so, so that each time I recall a memory I can model the experiences in the past event as the experiences of my current Self-model – and, therefore, "my" memories. An Episodic self is really just a model of me-as-other treated as a Self-model; but, by being treated as a Self-model, it, too, can become the source of pathologies of the self.

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

The next self is the most controversial of the selves: the **Narrative Self**. The nature of this self is a large part of the controversy: is the Narrative Self the life-story without which we cannot be a properly-functioning human? Or is it no more than a series of episodic selves which we treat as a continuous story when they do not need to be (Dennett, 1991)? Some people say that they do not have a Narrative Self, and we should not dismiss their feelings about their own selves as self-delusion (Strawson, 2004). However, even if the Narrative Self is not actually instantiated in the brain, it still works as a metaphor of a continuous self, and that may be enough for it to merit its place in SSH.

The Cultural Self: the self I should be

Like the Social Selves, the sixth self, the **Cultural Self**, is a model offered to the individual by others; but, unlike the Social Selves, it is not directly a model of the individual. Instead, it is a model of an ideal individual in this particular culture, explicitly the ideal persona that the addressed self should aim to be. As a marker of the expectations of others about me, I can usefully use any received Cultural Self as a scorecard for my social status. However, if it is treated uncritically as a target to be attained, the Cultural Self can become the cause of much misery for the modelling individual. The demand that a person adopt arbitrary roles or positions, contrary to their current Self-model (or, even worse, contrary to their Actual Self) is yet another source for pathologies of the self.

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

The final self, the **Projected Self**, is the self-image I want others to believe is me. It completes the circle of selfhood, which starts in the minds of others and, with the Projected Self, ends up back in the minds of others. It 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 the Cultural Self.

The Projected Self is an amalgam of my internal representations of myself and the expectations that others put upon me. Dörnyei (2009) refers to this target self, generated by internal as well as external expectations, as the Ideal Self. It is also the start of the next iteration of received Social Selves, because that big grey arrow pointing outward to the Universe links back to the people who generate the big orange inward arrows, which are the external inputs to the internal self-modelling process. Human selfhood is not an equation with an answer, nor is it a strange substance that we

can call upon as needed. It is a cyclical and iterative process in which each cycle changes the inputs being processed; and the process is kept in motion by our continued intentional social calculus and our attentional sharing of it.

The Exogrammic Self?

One significant moderator tradtionalof self-modelling is not addressed in the SSH: the effect of offline storage. The sharing of our individual social calculus systems allowed us to build bigger and more reliable social calculus systems, which require bigger and more efficient brains. Language, however, has given us a way to supplement our own brainpower when building our social calculus: we can store things outside our heads. I can tell things to other people in case I forget them, or merely to expand the knowledge pool; and I can also simply know that other people may know things I do not know, and ask them to fill in the gaps in my knowledge. These two processes of telling and asking can be used to preserve and extend my internal knowledge as external memories, exograms instead of engrams. Semiology and, later, writing gave us even more reliable exograms, but they also allowed other people a new way to share social models about us much more widely. Indeed, Bednarik (2014) suggests that the success of exograms may have been responsible for the puzzling 10% reduction in the size of Homo sapiens' brains during the last 40,000 years.

Nowadays, we deal with this Exogrammic Self without noticing that we are doing so: we manage our e-image daily as a matter of course, and an industry has grown up around this need. A Google search for "managing your e-profile" (with the quotes) produced 2.29 billion results, or one result for every three or four people on the planet. However, the need to address this Exogrammic Self is much more ancient than the Interweb: in 399 BCE, Socrates was called upon to defend himself against the charges that:

> "Socrates does criminal wrong (adikei) by not recognizing (ou nomizon) the gods that the state recognizes (nomizei), and furthermore by introducing new divinities (daimonia); and he also does criminal wrong by corrupting (diaphthairon) the youth (neous)." (Ober, 2016, p66.)

The charge of corrupting youth was for challenging their beliefs; so, essentially, all the charges were for philosophical corruption. This means that an exogrammic, shared social model of Socrates-the-criminal was put on trial and found guilty; but it was the actual Socrates who was executed. Even stranger, the actual Socrates colluded in his execution.

An Exogrammic Self may superficially seem to be just another Social Self received by the individual, but it has a developmental trajectory beyond the individual which makes it different from a Social Self. An Exogrammic Self is, like a Social Self, a third-person observation of an interaction between the individual and others which has been shared with the individual. However, the Exogrammic Self is not just a proxy for the person; the exogram achieves a level of accepted reality such that the person becomes a proxy for the exogram, and what was a reputational indicator becomes definitional. The Exogrammic Self is a major component of post-gatherer-hunter societies; but the SSH does not address it.

Why social calculus, selfhood and language may be linked

The SSH relies heavily on the idea that shared social calculus is the source of selfhood, and that language is the communicative tool that made sharing possible. However, a significant issue remains largely unaddressed: how were the three cognitive systems for social calculus, selfhood and language able to work together; what were the linking mechanisms? Recent research may have provided an answer: there may not be three systems working together, but one system doing three things.

It has long been known that the parietal cortex of the human brain is involved in planning of movement, mapping of locations in physical space, language construction and deconstruction, episodic memory, and declarative memory. However, several recent studies have indicated that it also has a coordinating role across a range of cognitions. For instance:

1. Thiery et al (2020) provide evidence that the inferior parietal cortex (IPC) in humans is involved in the exercise of choice between competing alternative strategies. This implicates the parietal cortex in modelling and choosing, key features of language and social calculus.

2. Humphreys et al (2021) propose a unifying model to account for the discovery that several cognitive tasks activate the same locations across the angular gyrus. They found that the mapping of aspects of events – semantic, social, spatial and temporal – all seem to use the same neurological resources, implying that they may operate using the same structures, rules and syntax.

3. Panichello & Buschman (2021) trained rhesus monkeys to switch between attention to a single stimulus and selection from a range of stimuli, and found that, in both attention and selection, the prefrontal cortex seems to act as a domain-general controller, allowing both attention and selection to be remapped to the IPC as standard templates. This implies that the coordinating role of the IPC may be a product of standardisation templates in the frontal cortex.

4. Numssen et al (2021) show that the IPC is involved in language, empathy, and attention, despite their different cognitive roles and products. They also show that there is high connectivity between posterior IPC and other parts of the cortex, and there is also lateralized functionality between the left and right hemisphere IPC areas.

Taken together, these papers indicate that the SSH is not directly refuted by current neurological approaches to social calculus, selfhood and language. However, it remains too early in our understanding of human cognition to treat these new findings as definitive, and thus move from hypothesis to theorem.

The role of the unconscious?

Now we come to the link between this presentation and the rest of the Open Foundation programme: the role of the unconscious. This is, for me, a stumble into an unmapped minefield; but, using the SSH as a guide, I would suggest the following may be significant for understanding the

unconscious will, and the way that pathologies of the unconscious might be therapeutically adjusted.

Traditionally, therapeutic interventions to selfhood have been directed at the inputs to the Social Self (the big orange arrows). These intervention routes are: the individual's perception of the opinions of others as expressed through the Social and Cultural Selves; and the individual's memory system as expressed through the Episodic, Narrative and Projected Selves. These interventions are not attempts to directly adjust the Self-model, instead they target the inputs to the Self-model; which means they are indirect solutions, and their efficacy is therefore variable.

However, there is one input to the Self-model I have not explored so far in this presentation, mainly because it cannot be consciously explored by the self. This is the input from the unknowable Actual Self to the Self-model. To explore this, the Actual Self has to become indirectly knowable in some way by the Self-model; and I believe that meditation, mindfulness, and even properly supervised psychoactive substances have an important role to play here. They provide a link between the Actual Self and the Self-model by making unconscious thoughts visible to the conscious as metaphors of conscious thoughts. While the Actual Self does not have consciousness, which makes it directly unknowable to conscious cognition, metaphorical representation can make it indirectly interpretable. This is important because the Actual Self remains the source for much of our selfness cogitation; and, while there is no agreement on how much of our cognition is subconscious, there is general agreement that less than half (possibly much less than half) is conscious. Therefore, accessing and interpreting subconscious cognition about self remains, in terms of our neurological understanding of human cognition, an important problem to solve.

Conclusion

This presentation has looked at two hypotheses. The first is about the evolution of selfhood, and the second is about the nature of selfhood. As hypotheses they should be internally consistent and consistent with the evidence presented; and I hope that you find them so. However, they are not theories: they have not been evidentially tested to an acceptable scientific standard. This presentation is therefore a starting point for further research; but then, so are all presentations.

References

Robert G. Bednarik (2014). Doing with less: Hominin brain atrophy. In *HOMO - Journal of Comparative Human Biology* 65, 433-449.

Aurèle Boussard, Adrian Fessel, Christina Oettmeier, Léa Briard, Hans-Günther Döbereiner & Audrey Dussutour (2021). Adaptive behaviour and learning in slime moulds: the role of oscillations. In *Philosophical Transactions of the Royal Society B* 376:20190757.

Brian Boyd (2017). The evolution of stories: from mimesis to language, from fact to fiction. In *WIREs Cognitive Science*, e1444.

Jonathon Brown (1997). Self-knowledge. In *The Self*. Psychology Press: Hove, E.Sussex, UK, ch4.

Eric M. Burton, Ibukun A. Akinyemi, Tiffany R. Frey, Huanzhou Xu, Xiaofan Li, Lai Jing Su, Jizu Zhi, Michael T. McIntosh & Sumita Bhaduri-McIntosh (2021). A heterochromatin inducing protein differentially recognizes self versus foreign genomes. In *PLoS Pathogens* 17:3, e1009447

Josep Call, Brian Hare, Malinda Carpenter & Michael Tomasello (2004). 'Unwilling' versus 'unable': chimpanzees' understanding of human intentional action. In *Developmental Science* 7:4, 488-498.

Daniel C. Dennett (1991). *Consciousness Explained*. Penguin Books: London, UK, pp413-414.

Zoltan Dörnyei (2009). The L2 Motivational Self System. In Zoltan Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*. Multilingual Matters: Bristol, UK, 9-42.

Robin I.M. Dunbar (1996). *Grooming, Gossip and the Evolution of Language*. Faber & Faber Ltd: London, UK.

Robin I.M. Dunbar (2004). *The Human Story: a new history of mankind's evolution*. Faber & Faber Ltd: London, UK, ch3.

Martin P.J. Edwardes (2019). *The Origins of Self: An anthropological perspective*. UCL Press: London, UK.

Sigmund Freud (1959[1926]). The Question of Lay Analysis. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XX (1925-1926): An Autobiographical Study, Inhibitions, Symptoms and Anxiety*. The Hogarth Press and the Institute of Psycho-analysis: London, UK.

Stephen J. Gould & Richard C. Lewontin (1997 [1979]). The Spandrels of San Marco and the Panglossian Paradigm: a critique of the adaptationist programme. In M. Ridley (ed.), *Evolution*. Oxford University Press: Oxford, UK.

Lauren Gravitz (2019). The importance of forgetting. In *Nature* 571, S12-S14.

Gina F. Humphreys, Matthew A. Lambon Ralph & Jon S. Simons (2021). A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition. In *Trends in Neurosciences* 1691.

Ole Numssen, Danilo Bzdok & Gesa Hartwigsen (2021). Functional specialization within the inferior parietal lobes across cognitive domains. In *eLife* 10:e63591.

Josiah Ober (2016). The Trial of Socrates as a Political Trial: Explaining 399 BCE. In Jens Meierhenrich & Devin O. Pendas (eds.), *Political Trials: Interdisciplinary Perspectives*. Cambridge University Press: Cambridge, UK, 65-87.

Monikaben Padariya, Umesh Kalathiya, Sara Mikac, Katarzyna Dziubek, Maria C. Tovar Fernandez, Ewa Sroka, Robin Fahraeus & Alicja Sznarkowska (2021). Viruses, cancer and non-self recognition. In *Open Biology* 11:200348.

Matthew F. Panichello & Timothy J. Buschman (2021). Shared mechanisms underlie the control of working memory and attention. In *Nature* online, https://doi.org/10.1038/s41586-021-03390-w.

Galen Strawson (2004). Not every life is a narrative: A fallacy of our age. In *Times Literary Supplement*, 15 October 2004, 13-15.

Thomas Thiery, Anne-Lise Saive, Etienne Combrisson, Arthur Dehgan, Julien Bastin, Philippe Kahane, Alain Berthoz, Jean-Philippe Lachaux & Karim Jerbi (2020). Decoding the neural dynamics of free choice in humans. In *PLoS Biology* 18:12, e3000864.

Charles Whitehead (2001). Social Mirrors and Shared Experiential Worlds. In *Journal of Consciousness Studies* 8:4, 3-36.

The Origins of Self: An anthropological perspective

Published by UCL Press.

You can download a free pdf copy at:

https://www.uclpress.co.uk/products/126003

(Paper copies, and an online HTML version, are also available.)