The biology of the imagination[i]

Simon Baron-Cohen

In what sense might something as intrinsically human as the imagination be biological? How could the products of the imagination – a novel, a painting, a sonata, a theory – be thought of as the result of biological matter? After all, such artefacts are what *culture* is made of. So why invoke biology? In this essay, I will argue that the *content* of the imagination is of course determined more by culture than biology. But the *capacity* to imagine owes more to biology than culture.

Let's start with a few definitional issues. What do we mean by 'imagination'? I do not mean mere imagery, though clearly the imagination may depend on the manipulation of imagery. Imagery is usually the product of one of the five senses (though it can also be generated without any sensory input at all, from the mere act of thinking or dreaming). Imagery typically comprises a *mental representation* of a state of affairs in the outside, physical world. I don't want to put you off from reading this essay by littering it with jargon, so let's just think of a mental representation as a picture in your head. That is what we are going to be calling an image, but that is not the same as imagination. Consider why not.

When we create a visual image of a specific object in our mind, the image as a picture of the object has a more or less *truthful* relationship to that object or outside state of affairs. If the image is a good, faithful, representation, it depicts the object or state of affairs accurately in all its detail. So, mental images typically have 'truth relationships' to the outside world. Of course, to create imagery in the first place depends on having the

relevant ‘hardware’. To create a photo, one needs a camera. To create a mental image, one needs a sense organ hooked up to a brain. An eye can do the trick, since the retina contains receptors that can code both position and colour in sufficient detail for the brain to which it is hooked up to create an accurate image. But in the absence of an eye, clearly an ear or a finger can do the trick too. With your ear, you can create an image of where that owl might be. With your finger, you can create an image of where your car-keys are.

Imagery may be necessary for human imagination. It has been suggested that all the products of the imagination are derived from imagery, following some transformation of the basic imagery. For example, Rutgers’ psychologist Alan Leslie, when he worked in London in the 1980s, proposed that imagination essentially involves three steps: Take what he called a ‘primary’ representation (which, as we have already established, is an image that has truth relations to the outside world). Then make a *copy* of this primary representation (Leslie calls this copy a ‘second-order’ representation). Finally, one can then introduce some *change* to this second-order representation, playing with its truth relationships to the outside world without jeopardising the important truth relationships that the original, primary representation needs to preserve. For Leslie, when you use your imagination, you leave your primary representation untouched (for important evolutionary reasons that we will come onto), but once you have a photocopy of this (as it were), you can do pretty much anything you like with it.[ii]

Let’s make this more concrete. Your eye looks at a fish. This causes your brain to form a visual image of a fish. So far, your primary representation ‘fish’ still has accurate truth relations with the outside world. The real fish has fins, eyes and gills, and so does your image of the fish. Or your eye looks at a woman, and this causes your brain to form

a visual image of the woman. Now you not only have a primary representation of a fish, but you also have a primary representation of a woman. This image, like the one of the fish, is also *truthful*. The woman you looked at has long hair and an alluring smile, and so does your primary representation of the woman.

In Leslie's important theory, to create such images or primary representations, the only hardware needed is a visual system that starts with an eye and ends in the visual cortex of the brain. But recall that that is only the first of his three steps. To move beyond imagery to imagination, to progress to steps two and three, one now needs an extra, special neurological mechanism. This extra mechanism can take each of the two primary representations (fish, and woman), and make *copies* of them. Whereas our brain previously just had two primary representations, it now has two second-order representations as well. So that was step two accomplished.

Finally, enter step three. This same special mechanism can now introduce *modifications* to the second-order representations at whim. It can for example *delete* some features on each of these second-order representations. Let's delete the head of the fish and delete the legs of the woman. And whilst we're at it, let's delete her long hair. Clearly these second-order representations are no longer *veridical*, that is, they no longer refer to anything in the outside world truthfully. But that's precisely the point. The brain is there as an evolved organ to represent what is going on in the outside world veridically. If there's a lion out there, the brain needs to know the image created by the visual system is accurate, so it can take the necessary action (fight or flight). But the human brain (whilst not wishing to sacrifice this important survival function of imagery) can be ratcheted up to do more than just represent the outside world veridically, and modifying

second-order representations opens up a world of new possibilities. It allows the brain to think about the possible, the hypothetical, about currently-untrue states of affairs.

Of course, *deleting* features from second-order representations are just the beginning of the set of possible changes that this mechanism can introduce. Another sort of change might be to *add* features to second-order representations that the primary representations from which they were derived never had. For example, adding snakes to the image of the woman. Or another kind of change this important mechanism can introduce is to *fuse* two second-order representations together. Just bolt them together to see what this would make. For example, the mechanism can combine the modified second-order representations, to produce the intriguing image of a woman with a fish' tail and with snakes coming out of her head. We can even give this newly formed second-order representation a name (mermaid).

Whereas any animal with a sense organ and a brain attached to it can produce an image (or a primary representation), there is a lively debate about whether any animals other than humans can produce second-order representations.[iii] Alan Leslie called the mechanism that can do steps two and three the 'meta-representational capacity' and he argued persuasively that this mechanism lies at the heart of the development of pretend play, and the human ability to mind-read.

Regarding pretend play, it has long been recognized that human infants from age 9-14 months old, begin to pretend. For example, they may pretend an object has features it does not have (e.g., pretending a toy tea cup is hot). Notice what is going on here. The infant brain has *added* a feature to the representation of the object that the object does not in reality possess. Or the infant may pretend the object has an identity it does not have

(e.g., pretending a toy tea cup has liquid in it). Or they may pretend one object is another (e.g. pretending a toy brick is a tea-cup, seen when the child puts the brick to a doll's lips, as if to offer her a drink). Such 'object-substitution', or playful manipulation of an object's features or identity, can take place *safely* if these modifications are made to second-order representations.

'Safely' in what sense? In the sense that the developing infant brain needs to keep track of what objects are really like in the real world. The brain needs to be able to distinguish between representations of objects that have some claim to be truthful (my eye tells my brain this object is a fish), from those representations that have no claim to being true (I imagine a creature called a mermaid). If the infant brain was introducing such modifications to the primary representations themselves, they would no longer be able to sort out what was real and what was not. This could lead to the infant ending up seriously confused or even deluded about the nature of objects (do fish have women's heads?). It could also lead the brain to fail to distinguish a real threat (this is a lion) from an imagined threat (this is a pretend lion). The brain has paid for itself in evolution, not by wreaking havoc with the veracity of primary representations, but by quarantining the truth relationships of primary representations.

Primary representations have the evolutionary function of representing the world faithfully, in order to build up a knowledge base of what the world is *really* like. Change your primary representations and you risk jeopardising the quality and reliability of your knowledge base, your database of what reality consists of. Leslie's important insight was that we know the normal infant is not confused by pretend play. They do not for a moment believe the pretend tea-cup really is hot. They *know* that it is not, because their

primary representation (tea-cup = cold) has been left untouched. And this is only possible because of step two above. By making a *copy* of the primary representation this has, in Leslie's chilling phrase, 'quarantined' the truth. In the second-order representation, none of the usual truth relationships need apply. The pretend tea cup can be hot even whilst the real tea-cup is cold. The primary and second-order representations are divorced and can have different functions. The function of a second-order representation is to allow the brain to manipulate truth in an infinite number of ways, to explore possible rather than real states of affairs. Pretend play does not just allow you to play. It allows you to 'imagine' hypothetical worlds, arguably a prerequisite for the serious enterprise of planning and engineering, as well as art or science.

In what sense might a meta-representational capacity be essential for mind-reading? Let's define mind-reading as the ability to put yourself in someone else's shoes, to imagine the other person's thoughts and feelings.[iv] Leslie's deeply interesting argument is that when you mind-read, you again need to quarantine your primary representations. Here's how his argument goes. Just as your mental picture of a fish has 'truth relations' to a real fish in the outside world, so a belief, or a sentence, has truth relations to real events in the outside world. Thus, 'John is having an affair with his colleague' is a primary representation of a state of affairs, and is true if John is indeed having an affair with his colleague. But when we mind-read, we again take the primary representation (step one), *copy* it so that it becomes a second-order representation (step two), and can then *add* a prefix (step three) that completely change its truth relations with the outside world.

Thus, we can take the primary representation ‘John is having an affair with his colleague’ (step one). We can copy it to produce an identical version ‘John is having an affair with his colleague’, except this version is tagged as being a copy or a second-order representation (step two). Finally, we can add a prefix such as ‘Mary believes that’ to the second-order representation to end up with ‘Mary believes that “John is having an affair with his colleague” ’ (step three).

Such second-order representations have unique logical properties, an insight that Leslie borrowed from the standard views in philosophy of mind. They have, to use the jargon, referential opacity. ‘I pretend that “this tea-cup is hot” ’ is true if I pretend this, irrespective of whether the tea-cup really is hot. ‘Mary believes that “John is having an affair with his colleague” ’ is true if Mary believes it, irrespective of whether John really is having an affair. According to Leslie, and I think he is right, when we mind-read (just as when we use imagination), we employ such second-order representations. I can maintain my own knowledge base (John is not having an affair) whilst representing someone else’s different (possibly false) belief (Mary believes he is). I can maintain my own realistic, true perception of the outside world (this is a lion) whilst representing possible and imaginary creatures (a lion with two heads). To mind-read, or to imagine the world from someone else’s different perspective, one has to switch from one’s own primary representations (what one takes to be true of the world) to someone else’s representation (what *they* take to be true of the world, even if this could be untrue). Arguably, empathy, dialogue, and relationships are all impossible without such an ability to switch between our primary and our second-order representations.

So, what has all this got to do with the original question of whether the capacity for human imagination is, at its core, biological? For Leslie, the capacity for meta-representation involves a special module in the brain, which humans have and that possibly no other species possesses. In the vast majority of the population, this module functions well. It can be seen in the normal infant at 14 months old who can introduce pretence into their play; seen in the normal 4 year old child who can employ mind-reading in their relationships and thus appreciate different points of view; or seen in the adult novelist who can imagine all sorts of scenarios that exist nowhere except in her own imagination, and in the imagination of her reader.

But sometimes this module can fail to develop in the normal way. A child might be *delayed* in developing this special piece of hardware: meta-representation. The consequence would be that they find it hard to mind-read others. This appears to be the case in children with Asperger Syndrome. They have degrees of difficulty with mind-reading.[v] Or they may never develop meta-representation, such that they are effectively 'mind-blind'. This appears to be the case in children with severe or extreme (classic) autism. Given that classic autism and Asperger Syndrome are both sub-groups on what is today recognized as the 'autistic spectrum', and that this spectrum appears to be caused by *genetic* factors affecting brain development, the inference from this is that the capacity for meta-representation itself may depend on genes that can build the relevant brain structures, that allow us to imagine other people's worlds.

What are the consequences for people on the autistic spectrum, and for our understanding of the role of biology in human imagination? Children with severe or classic autism may end up with an exclusive interest in the real world, with no interest at

all either in mind-reading, pretending, or fiction. They may enjoy making patterns with real objects, or watching how real objects behave, but not even spare a thought for how someone else might be feeling or what they might be thinking, or understand why a mermaid or a unicorn is a fun idea. Children with Asperger Syndrome may manage to mind-read to some extent, after a delay in developing this skill. But their delay may mean they still find empathy challenging even in adulthood. They may show a preference for factual reading material over fiction, or for documentaries over fictional films, perhaps because the hardware in their brain that functions to form primary representations and understand the real world of physical objects is more highly developed than the meta-representational hardware in their brain that functions to represent possible states of mind.

Since the disability that comprises classic autism is biological in origin, then children with autism are offering us a big clue about the biological basis of the imagination. Of course, when the meta-representational hardware develops normally, biology has done its job. From then on, the *content* of our imagination, whether we imagine an angry god or a school of wizardry, a mermaid or a devil, owes more to our specific culture than to biology. But the capacity to imagine depends on genes that build brains with a very specific kind of mechanism – one that we take for granted whenever we form relationships or fantasize.

Notes

[i] I am grateful for the support of the Medical Research Council, UK. This essay is reprinted from an essay in Headlem Wells, R & McFadden, J (eds) What is human nature? The Continuum International Publishing Group, London.

[ii] A.M. Leslie, 'Pretense and Representation: The Origins of "Theory of Mind" ', *Psychological Review*, 94 (1987), pp. 412–426.

[iii] Josef Perner, *Understanding the Representational Mind* (Cambridge, Mass.: MIT Press, 1991).

[iv] See Simon Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (Cambridge, Mass. and London: MIT Press, 1995).

[v] See Simon Baron-Cohen, *The Essential Difference: Men, Women and the Extreme Male Brain* (London: Allen Lane, 2003).