JEAN-CHARLES PELLAND
New College of the Humanities;
McGill University
jeancharles.pelland@mail.mcgill.ca

GROUNDING NORMATIVITY IN BIOLOGY: THE UNEXPRESSED RULES OF CORE COGNITION[1]

abstract

Saul Kripke's (1982) sceptical take on Wittgenstein's rule-following paradox challenges us to find facts that can justify one interpretation of a symbol's past use over another. While Ruth Millikan (1990) has answered this challenge by appealing to biological purposes, her answer has been criticized for failing to account for the normativity of rules like addition, which require explicit representations. In this paper, I offer a defense of Millikan. I claim that we can explain how we build intentions to add from the content of core cognition modules like the approximate number system, and argue that Millikan's answer is better equipped to explain the origins of rules than communitarian approaches like that endorsed by Kusch (2006). I then explore the worth of pluralism about rules and try to find common ground between expressed and unexpressed rules in terms of expectations of how the world is supposed to behave.

keywords

Kripkenstein, biosemantics, normativity, biological functions, core cognition, numerical cognition, approximate number system, rule-following

1 Thanks to Brian Ball for commenting on previous versions of this paper and for many helpful discussions on this topic. Thanks also to Alain Voizard. Work on this paper was made possible by Grant 207944 from the Fonds de Recherche du Québec – Société et Culture.

Phenomenology and Mind, n. 17 - 2019, pp. 198-209
DOI: 10.13128/pam-8037
Web: www.fupress.net/index.php/pam

ISSN 2280-7853 (print) - ISSN 2239-4028 (on line)

1. Introduction

A common, seemingly platitudinous claim about meaning – at least, for words and symbols – is that it is somehow normative:[1] if a symbol has a particular meaning, then there are rules dictating correct and incorrect ways of using it. When using the symbol '+', for example, we tend to think that the meaning attached is the addition function, and that there are correct and incorrect ways of adding. Threatening our meaning-determinist inclinations, Saul Kripke's (1982) sceptical take on Wittgenstein's rule-following paradox challenges us to find facts that can justify one interpretation of a symbol's past use over another. After considering tempting answers like facts about dispositions and intentions, Kripke's sceptic claims no such facts can be found, which pushes him to opt for an interpretation of the meaning of symbols that is based not on metaphysical facts, but on social convention. This communitarian solution, while popular, has problems explaining the origins of social convention, which, on pain of regress, cannot be a matter of social convention itself. Ruth Millikan's (1990) naturalization of intentions in terms of biological purposes provides a promising solution to Wittgenstein's paradox, rescuing some of our intuitions about the objectivity and determinacy of language and meaning by appealing to biology. However, attempts to ground normativity in biology have come under fire. A number of authors complain that biological norms are not real norms (e.g. Kusch 2006, Hutto & Satne 2015).

This paper offers a response to these criticisms of biological rules. I start off by summarizing Kripke's challenge in section 2, before presenting some of Millikan's response to this in section 3. In section 4, I argue that we can answer Kripke's challenge by applying Millikan's biological purposes to recent work in numerical cognition, allowing us to explain how we build intentions to add from the content of core cognition modules like the approximate number system. This is followed in section 5 by a response to Martin Kusch's (2006) criticism of Millikan. Then, in section 6, I try to reconcile Millikan with her critics by adopting pluralism about rules. In section 7 I end by trying to find common ground between expressed and unexpressed rules in terms of expectations of how the world is supposed to behave.

1 See e.g., Whiting 2007 for a defence of this claim. See also Whiting 2013 for a comparison of competing accounts of the normativity of meaning.

2. Kripke's Wittgenstein

Observing that our past use of a symbol can be described using an infinite number of rules, Wittgenstein famously wrote: "this was our paradox: no course of action could be determined by a rule, because every course of action can be made to accord with the rule." (Wittgenstein, 2001, §201) Building on Wittgenstein's observation, Kripke imagines a sceptic who challenges us to come up with a fact that allows us to determine that someone means addition by '+' rather than 'quaddition', where quaddition defines the quus function (⊕) as follows:
$x \oplus y = x + y$, if $x, y < 57$; $x \oplus y = 5$ otherwise.[2]
Kripke argues that any fact about the past use of a symbol – be it the intention to use it in a certain way, the qualia associated with its use, the dispositions to use it in various circumstances, etc. – is consistent with any number of quaddition-like rules. If Kripke is right, this implies that there are no facts about whether we meant addition rather than quaddition when we used '+'. Since Kripke's challenge needn't be restricted to '+', nor even to symbol use, the problem quickly spreads to *any* rule, linguistic or not, forcing us to wonder how there could there be facts about whether we are following one rule rather than another. Kripke's sceptical challenge, if left unanswered, leaves no room for determinate rule-following of any kind.
A tempting way to answer Kripke's sceptic is by appealing to intentions: when I use the symbol '+', I mean addition, not quaddition, because in the past, I intended to add things, not quadd things. Unfortunately, as Wittgenstein's discussion of the regress of explanations taught us (e.g., in 2001: §85-88), when we try to identify what constitutes an intention to follow a rule – say, addition – it seems the only answer we can give involves intentions to follow other rules. But then this more basic intention is just a 'rule for interpreting a rule', to which the sceptic can pose his challenge anew. For example, if I wanted to say that my meaning addition by '+' consisted in my intention to count both addends, the sceptic could simply ask what my intention to count consisted in, and how it differed from, say, quounting, where quounting is identical to counting for all past behaviour, but diverges after a yet-to-be counted (or quounted) number.
Answering the sceptic by appealing to explicit representations of rules like intentions leads us to regress, so it looks like we must abandon intentions as a potential solution to the challenge. And yet, this only follows if the only way to explain what it means to intend to follow a rule is by referring to another intention. One way to answer the sceptic, then, would be to explain what intentions to follow rules are *without* recourse to other explicit representations. Ruth Millikan's (1990) naturalization of intentions allows us to do just that.

3. Millikan's straight answer: biological norms

Millikan delivers a biological approach to rules in which norms for biological entities can be derived from their biological purposes. Biological purposes are rules set by evolutionary design to which entities may or may not conform – more often than not, without being aware of it. The basic idea is that even though biological entities are disposed to behave in many ways, only some of this behaviour accounts for the proliferation of the entity's ancestors. Those dispositions responsible for a behaviour that explains why they have survived in the past are singled out as being in accordance with the norm set by the entities' biological purposes.[3]
On this view, to assess whether an entity's behaviour is correct, we need to refer to its history and look for behaviour and dispositions that have allowed its species to thrive:

2 Needless to say, the restriction to numbers smaller than 57 is merely meant to simplify the discussion. Basically, quaddition is a function identical to addition for all sums that have already been calculated, but whose value differs after a certain yet-to-be-calculated pair of numbers.
3 This is, of course, an overly simplified summary of Millikan's account of biological purposes. For a more elaborate discussion, see Millikan 1984 (esp. Chs. 1 and 2), as well as Millikan 1989a, 1989b. See also Shea 2006.

> it is the reference to evolutionary history that [does] all of the work in explaining how norms come to apply to the activities of an animal, in explaining how there can be a standard from which the facts of individual behaviour diverge. (Millikan, 1990, p. 337)

Simple examples of biological purposes include the blood-pumping function of our hearts, the communication function of bees' mating dances, and the reproductive purpose of interceptive flight patterns of male hoverflies. At a more complex level we find biological functions that drive animal learning. Such learning is a biologically-determined rule that governs the behaviour of the animals' nervous systems and has helped their species thrive throughout history by allowing them to adapt to their environment – which, in some cases, includes complex social networks. For example, by rewiring the neural connections between smells and memories, such learning mechanisms allow rats to avoid eating food that previously made them sick, thus decreasing their chances of getting poisoned.

Applying her theory to the complex biological purposes of humans, Millikan proposes that we think of human intentions as a kind of biological purpose. Much like we can explain a circus poodle's ability to learn how to ride a tricycle in terms of the evolutionary advantage conferred by its innate learning mechanisms, we can describe the mechanisms and dispositions behind intentions to follow rules like addition (or quaddition) in terms of evolutionarily-inherited biological purposes. Like the poodle's cycling abilities, the human ability to form intentions is grounded in neural mechanisms that need not be driven by any explicit representation or intention: "the unexpressed purposes that lie behind acts of explicit purposing are biological purposes." (Millikan, 1990, p.330). On this view, my intention to add is the result of how my experience has shaped innate biological mechanisms handed down by evolutionary design.

How can this help us answer the sceptic? By putting an end to the regress of intentions described above:

> Explicitly meaning or intending, if this requires representing what one intends, presupposes a prior purposing: purposing to let the representation guide one in a certain way...this prior purposing cannot be analysed as the original explicit purposing was analysed without regress. Rather, a prior unexpressed purposing must be assumed. The reasonable conclusion seems to be that ordinary explicit intending rests on biological purposing – biologically purposing to be guided by, to react this way rather than that to, one's representations. (Millikan, 1990, pp. 342-3)

Millikan's classification of intentions as biological purposes gives us a framework in which we can express facts about intentional behaviour without needing to appeal to further intentions. Rather, to describe the content of an intention, we need to look at an individual's history, environment, and neurological makeup. In so doing, we do not enter a regress – at least, not an unending regress of intentions, since, in the end, some representations guide biological entities without the entity being aware of it.

For example, consider the 'hoverfly rule': if a male hoverfly detects female-like stimuli, it enters an interceptive flight pattern, guided by representations and mechanisms of which it is (presumably) not aware. Similarly, while some human intentions may owe their content to other intentions, Millikan's idea is that some intentions are formed from representations that guide our behaviour without our being aware of their content. To illustrate how this might help us answer Kripke's sceptic, in the next section, I discuss how we could exploit recent work in core cognition to explain how we form the intention to add from evolutionarily-inherited content, which, at least initially, we need not be able to form any intentions about.

4. Core cognition and addition

Here, we want to use Millikan's biological purposes to stop a regress of intentions in answering the sceptic. Given that intentions to add develop out of previous intentions – including intentions to count – this means that at some point not too far down the explanatory ladder we need to find facts about non-explicit representations involved in forming intentions to count, so that we can explain how these differ from intentions to quount. But what sort of non-explicit representations guide us towards counting? To answer this, it may be useful to look at recent progress made in the study of how we think about numbers.

Thanks to methodological and technological advances in the study of numerical cognition, many theories of the development of representations with numerical content have sprung up in recent years (e.g. Dehaene 2011; Carey 2009; DeCruz 2008; Leibovich et al. 2017; Cohen Kadosh & Walsh 2009). Research in this burgeoning field has yielded tremendous findings, including the discovery of innate cognitive systems that may serve as the building blocks of our formal arithmetical abilities. While there is still considerable debate regarding how these systems interact and develop to allow us to form mathematically-viable numerical content, there is almost universal agreement that at least two cognitive systems – the so-called approximate number system (ANS) and the object-file system (OFS) – supply the basic content that eventually allows the construction of explicit numerical representations. In a nutshell, the ANS tells us how many things there are in a part of the environment we are paying attention to, but its precision decreases as the number of things it tracks increases. As for the OFS, while it is not dedicated to tracking numbers of items, it does allow us to track the spatiotemporal properties of up to four objects at a time, and can respond to changes of numbers of items within this restricted numerical range.[4]

The important point to consider here is that there is ample evidence from habituation and violation-of-expectation studies involving animals and pre-verbal infants that these systems are innate and evolutionarily-ancient. This means they produce representations with quantity-related content in organisms whose ability to form explicit intentions is highly doubtful, if at all plausible. For example, there is little reason to think that rats can form intentions, and yet they are equipped with groups of neurons whose unexpressed biological purpose is to take input from the senses to yield non-explicit representations with quantitative content (Meck & Church 1983). Core cognition modules like the ANS and the OFS are often seen as supplying building blocks for more advanced cognitive abilities like understanding what agents, objects, and numbers are (Carey 2009), as well as providing a basis on which many features of language, such as the mass/count distinction, are built (Strickland, 2016; Odic 2014).[5] Although details about how we build complex concepts from core cognition modules have not been ironed out, the strategy is usually to appeal to the effects of more general cognitive faculties like associative learning, general learning mechanisms like induction, or language.

How can this help us with the sceptic? By illustrating how intentions to add could be built from representations that can guide our behaviour unintentionally, like those produced by the ANS and the OFS and the learning mechanisms that operate on these. When Kripke's sceptic asks for a fact that can allow us to determine that we are adding rather than quadding, we can appeal to facts about the content of core cognition modules to stop the regress of intentions when explaining why we are adders, not quadders. This is because such innate systems wear

4 While the literature on these systems has grown exponentially in recent years, an easy and short introduction to them remains Feigenson et al. 2004. For a treasure trove of findings and interpretations, see Cohen Kadosh & Dowker 2015.

5 This being said, see Rips 2017 for a sceptical take on such attempts to explain conceptual development from core cognition.

their rules on their sleeve, so to speak: their content, which has been carved by Darwinian processes, is forced upon us. If the rules that govern these systems are not amenable to Kripke's sceptic, then if we build our explicit rule-following from them, then we can expect our explicit rule-following to be quus-free.

As it turns out, there is good reason to think that unexpressed rules like those that govern how innate systems like the ANS behave can't be quussed. First, as sketched above, these rules have been set by evolutionary history. This biological standard means that any quus-like reintepretation of rules derived from biological functions are simply less efficient descriptions of why certain traits have helped organisms survive. For example, while we could potentially explain hoverfly behavior in terms of a quoverfly rule, the active part of the explanation would be solely done by the part of this rule that coincides with the hoverfly rule, while the Kripkean part of the rule does no explaining whatsoever, as Millikan points out. Indeed, any number of quus-like rules could describe the behaviour of systems like the ANS, but only the non-quussed aspect of such rules explains their role in helping organisms equipped with the systems thrive. But "a complexity that can simply be dropped from the explanans without affecting the tightness of the relation of explanans to explanandum is not a *functioning* part of the explanation." (Millikan, 1990, p. 334) Further, and perhaps more importantly, quus-like reinterpretations are in a sense parasitic on their regular counterparts: for example, one cannot build a quus rule without first having a plus rule.

So, keeping in mind that the unexpressed rules followed by biological systems are immune from Kripke's sceptic, it looks like we have a story to tell about how intentions with numerical content, such as intending to count and intending to add, are grounded in the biological purposes of neural systems tuned to quantity-related information in the environment, like the ANS. I intend to use the symbol '+' to mean addition because I learned to associate this symbol with the act of adding, not quadding. I am adding, not quadding, because when learning to add I learned how to count the addends, not quount them, and my intentions to add are built from being guided by this counting-related content. When pressed further by the sceptic, this time regarding intentions to count, not quount, I can appeal to the content used to learn how to count: when learning to count, I learned how to associate number words to representations of discrete numerical quantities, and counting involves being guided by this association.[6]

Now things get interesting: if the sceptic tries to question numerical quantities, perhaps with numerical quuantities, I can appeal to the content produced by core cognition modules to stop the regress of intentions. This is because the rule I follow when being guided by representations of quantities in counting is not a rule I set for myself intentionally via a learning process. Rather, the rule I follow when entertaining content produced by core cognition modules like the ANS and the OFS is a biological rule which I have no choice to follow. The same can be said for the rules I follow when being guided by general learning mechanisms like induction and associative learning. My intention to count is built from these innate systems, whose rules are unexpressed and unquussed. This means we have a rough sketch of how we can give facts about intentions to add without running the risk of running in endless explanatory circles.

The upshot is that intentions are fair game in answering the sceptic, and we can meet Kripke's challenge by providing biological facts about where intentions come from.

6 As mentioned above, the details of how we build precise numerical content from the content of systems like the ANS have yet to be settled, but this need not prevent us from imagining how such systems are involved in learning to count. For example, there is strong evidence that the words in our count list eventually get mapped onto the content of systems like the ANS (Dehaene 2011, Dehaene & Cohen 2007), though as Carey (2009) points out, there is evidence that this mapping occurs *after* we have mastered the use of number words beyond the subitizing range (i.e. larger than four).

5. Biological rules?

Setting aside potential issues concerning the accuracy of this sketch, there is an important question raised by saying that expressed rules like addition and unexpressed biological rules like those that apply to the functioning of systems like the ANS and the OFS are both instances of rule-following. The question concerns the relation between the unexpressed content behind biological rules like those that apply to hoverflies and the ANS and the expressed content involved in intentions to follow conventions like those involved in arithmetic. The relation here needs to be characterized in a way such that we can read off the normativity we associate with intentional behaviour with the normativity we find in biological functions. But biological and social norms seem to be of a different kind, since only the latter rely on understanding or explicitly representing a rule. In social rule-following, the fact that we are voluntarily and consciously being guided by explicitly represented content somehow seems to make a difference to the legitimacy of the rule: only rules that can be explicitly, voluntarily learned are supposed to count as legitimate rules.

This issue has been expressed in a variety of ways over the years.[7] For example, Martin Kusch (2006) expresses his doubts about whether Millikan's biological purposes capture the literal sense of normativity thus:

> Using normative language to describe proper functions seems inadequate. The norms and standards in question do not literally have a norm-authority, that is, someone who has introduced them and who sanctions deviation. How then are we to think of talk of biological norms? Is this talk not merely metaphorical? If so, then it is hard to accept that Millikan has given us a naturalization of meaning and normativity. After all, in order to understand the metaphorical sense of normativity we first have to understand the literal sense. And this literal sense presupposes an understanding of (expressed) intentions. (Kusch, 2006, p.73)

Here, Kusch proceeds on the assumption that, to be considered legitimate, a norm must be introduced and sanctioned by a 'norm-authority'. The claim is that since there are no such norm-authorities behind evolutionary design, biological purposes cannot be considered legitimate setters of norms. As I hope to show in this section, it is difficult to see how to frame this objection to biological normativity on solid grounds.

The main problem here is that Kusch begs the question of the origin of normativity by requiring that norms be (intentionally) introduced by norm-authorities. Kusch's claim that the normativity of biological purposes is a metaphorical application of legitimate normativity and that the literal sense of normativity requires understanding intentions is problematic for the same reason.[8] For while it could turn out that we do indeed need norm-authorities and intentions to have real norms, simply positing this without arguing for it is not enough to dismiss the possibility that normativity be grounded in biology.

On the contrary, taking a closer look at these norm-authorities raises a few difficult questions, some of which might be answered by appealing to biological functions. For example, how could a norm authority be able to go through the infinite uses of a word and 'sanction

7 See Hutto & Satne's 2015 discussion of Neo-Cartesianism for a few examples.

8 Kusch's rejection of Millikan's proposed grounding of normativity in biology is reminiscent of Bloor's, who claims that rule-following is an *actor's* category, and that an 'actor's own awareness of these norms is constitutive of their very existence as norms' (Bloor, 1997, p.105) And yet, children learning how to dance or other forms of learning by imitation do not seem to require knowledge above anything like 'do what others are doing', and it is not obvious that every case of learning by imitation involves explicit representation of the rule being learned. Cases of animals learning how to ride tricycles or tie knots seem like obvious examples of this.

deviation' in advance? It seems that answering this question would require appeal to some kind of rule for sanctioning deviation, but this rule would require further norm-authorities and further sanctioning, thus entering us into a regress much like that discussed above regarding intentions to follow rules. In other words, if norm-authorities set norms, who sets the norm-authority's norms of behaviour? To answer this last question, it seems we either need to appeal to another norm-authority, thus generating a regress, or appeal to another mechanism that sets norms. If the latter option holds, then why bother with norm-authorities at all? A similar regress seems to threaten the popular communitarian answer to Kripkenstein: if normativity is a matter of agreement, then how can we avoid a regress when agreeing on what agreements are? Put in a Kripke-like way, if we all need to agree in order for a practice to become canonized as setting a norm, then how do we know we are agreeing with each other, rather than quagreeing?[9]

Moreover, if normativity is indeed grounded in intentional social agreement, as Kusch proposes, one wonders how these intentions get their normative force, if not from biological facts about how humans (and other animals) are hardwired to react preferentially to their conspecifics. After all, why would we be inclined to follow social rules at all unless we were biologically hardwired to do so? If anything, Kusch's norm-authorities seem to support – if not require – Millikan's grounding of normativity in biology, given that it remains a mystery where rules for their behaviour would come from, if not biology.

Consider for example rules of language: like those for biological organisms, many linguistic rules emerge as a result of Darwinian processes, because certain practices are better adapted to their environment than others, not because someone has conferred a preferential status upon them. It is difficult to see this cultural Darwinism as being exclusively the result of someone introducing rules and sanctioning deviation. There are many reasons that can explain why some practices become the norm while others fail.[10] For example, when mobile technology became ubiquitous, countless new words and expressions emerged and rapidly spread worldwide. Abbreviations, symbols, short words, all these were better suited to small keyboards, screens, and time constraints. What did *not* happen is that someone said "and now I confer thee the normal use of LOL".

This being said, even though Kripke's communitarianism is premised on the fact that there is no biological story capable of satisfying the sceptic, as illustrated by his lengthy analysis of the limitations of dispositionalism, there is a sense in which this communitarian response is compatible with Millikan's biological dispositionalism. After all, even if social conventions in general are rooted in biological purposes, this need not imply that they don't cement the meaning of specific symbols like '+'. And yet, there is a sense in which the communitarian response cannot satisfy the sceptic unless it is supported by a story like Millikan's. As just discussed, the communitarian response is not equipped to answer the sceptic's probe into how we set agreements apart from quagreements, and leaves the origins of such social norms unanswered.

But perhaps more importantly, while the facts that can explain why I mean addition by '+' include my socio-cultural history and the agreements that have been made in the past on top of my biological ancestry and individual ontogeny, the communitarian answer leaves out the fact that symbols have meaning to individuals in virtue of being associated with mental content. In many cases, symbols force specific content upon us regardless of social rules and conventions. For example, things that look like faces – including drawings of faces, but also

9 For more on this, see the review of issues with Neo-Pragmatism by Hutto & Satne 2015.
10 See Richerson & Boyd 2005 for an account of such mechanisms of cultural evolution.

arrangements of dots like stars in the sky and knots in wooden planks – will elicit content about faces irrespective of social contexts, since humans (and some other animals) are born with the ability to detect conspecifics by giving priority to certain configurations of stimuli.[11] It is highly doubtful that any amount of social coaching could manage to override such biologically-driven content.

Further, the limitations of the communitarian answer become apparent when we consider those members of societies whose biological makeup is malfunctioning for one reason or another. If facts about the meaning of symbols were constituted solely by social conventions, then all we would need in order for individuals to learn the meaning of symbols would be to satisfy conditions for establishing social conventions. But many individuals that can enter into social and linguistic conventions nevertheless fail to manage to learn the meaning of symbols like '+' due to learning or brain deficits. A particularly relevant example here is dyscalculia, which can prevent of significantly curb the ability to process numerical information. In some cases, no amount of social coaching or training can allow a person to overcome the limitations that accompany such deficits and allow them to associate symbols with the intended mental content, simply because the individual is unable to entertain it.

This illustrates how important it is to keep in mind that there is an ineliminable biological component that underlies social agreements and that without the relevant mental content in individuals' heads, it would be impossible to agree on anything. Arguably, without species-specific modules like those of core cognition of agency in humans,[12] any 'social' behavior would be the equivalent of schooling fish, whose coordinated movements can present the illusion of explicit organization, even though no one would claim that fish need to agree in order to move in harmony. If this is true, then even though many rules could not be learned outside a scaffolded social context, claiming that the meaning of a symbol like '+' is constituted by social agreements would miss an important part of the story of why such symbols are meaningful to individuals.

6. How many types of rules?

It seems we are stuck in a difficult position: on the one hand, we saw that there are thorny problems associated with views like those offered by Kusch. On the other hand, despite these issues, it does feel like the normativity of explicitly represented rules is different from that of unexpressed rules. One way out could perhaps be to accept that there are many types of rules, and none of these has priority over the other. For example, Millikan points out that are many types of norms:

> By 'normative' philosophers typically have meant something prescriptive or evaluative, but there are other kinds of norms as well. There are non-evaluative measures from which the facts or from which instances can depart; for example, a simple average is also a kind of norm. I argue that the central norms applying to language are nonevaluative. They are much like those norms of function and behavior that account for the survival and proliferation of biological species. Broadly speaking, they are biological norms. (Millikan, 2005, p. vi)

11 Of course, this is not to say that all animals share the same innate content. While we can easily show that humans share much of their innate cognitive and perceptual toolkits with other animals, it is equally easy to show that there is tremendous variation between species on what sort of content gets preferential treatment. Indigo-buntings, for example, are born with an ability that is sadly lacking in humans, that of being able to process data from the night sky to effortlessly compute where the north lies (Emlen 1975).

12 See chapter 5 of Carey 2009.

Why would intentional behaviour somehow be more worthy of normative status, as Kusch and others have claimed? After all, Millikan's biological purposes seem to have much in common with Kusch's 'legitimate' norms. For example, both expressed and unexpressed purposes involve being guided by representations. This applies to hoverflies, but also, in radically different ways, to humans. Experiments involving subliminal priming are just one of many examples of representations guiding human behaviour without our being aware of it: in such cases, our senses pick up data that do not make it to consciousness (say, because they are presented too briefly) and yet they have a clear impact on our intentional behaviour. Is this behaviour not the direct result of unexpressed biological functions? If so, then why would we not be able to classify other behaviour resulting from unexpressed biological functions – including hoverfly flight patterns – as equally worthy of normative evaluation?

Another consideration that seems to support unexpressed purposes as legitimate yardsticks of normativity is that, in many cases, even for those purposes that are originally explicitly represented, it is possible to follow the same rule without intending to. For example, at some point, tying one's shoes becomes so ingrained in our routine that we need not be aware that we are doing it. In such cases, however, unexpressed purposes are parasitic on previous occurrences of actual intentions, like intending to imitate mommy's (shoe-tying) actions, which would still require us to attribute a different status to explicitly represented rules.

What about unexpressed purposes that do not result from internalized rules? For example, can we say that the hoverfly rule involves the same kind of normativity as rules for tying one's shoes? Here, the answer is not so clear. On the one hand, it is common to think of unusual behaviour as going against established rules, even when talking about the behaviour of flies and organs. For example, if a heart beats erratically, we think its behaviour isn't conforming to its biological purpose, and we look for a cure. If a hoverfly doesn't conform to its usual flight pattern, we will think it is injured or that it is windy outside. If a person tries to tie their shoes by setting them on fire, we will think she is not following shoe-tying rules. Compare these cases with, say, calling a rock's shape abnormal. Whereas the rock may indeed have a shape that is statistically unlike most rocks, calling its shape abnormal has nothing to do with the shape it *should* have. Perhaps, here, we have a metaphorical use of normative language, or we are using a different type of norm. But biological purposes are established by reference to etiological concerns, irrespective of statistics – in fact, many biological rules fail to be followed most of the time – and thus carry a different normative component – much like the one involved in explicitly represented purposes.

On the other hand, it is common to speak of a person understanding the rules of a game, or a person following rules of etiquette, but it is not common to speak of hoverfly rules or a heart's following biological rules (at least, this is not common outside of philosophical circles). There is an aspect to rule-following that does seem to imply explicitly representing the rule in question, and this aspect is absent in most of following the rules set by evolution. Even by accepting pluralism about rules, there is a lingering problem of how biological rules relate to explicit rules: what makes them all *rules*?

7. Rules and expectations

So, how does our ability to follow explicit rules like addition relate to our ability to follow non-explicit rules like those that describe the functioning of systems like the ANS? Perhaps an answer would be to accept that not all rules require conscious awareness to be followed, but that all rules involve being guided by representations of how the world is supposed to behave. We could then explain differences in types of rules by appealing to differences in representations of how the world is supposed to behave: when these expectations come solely from innate sources like genes or core cognition modules, we could talk of unexpressed rules. When expectations are the result of learning, we could talk of expressed rules.

This way, we can accept that hoverflies can follow rules, since the hoverfly rule, though not explicitly available to the hoverfly itself, does involve a representation of how the world is expected to behave – i.e., the world is supposed to contain female hoverflies and entering into a certain flight pattern is supposed to increase chances of successful mating. The mechanism allowing hoverflies to follow the hoverfly rule is built on other mechanisms that evolved because the world contained female hoverflies that could be intercepted in certain ways. Going up the cognitive ladder (and adding a few modules to it), we can frame the unexpressed rules involved in systems like the ANS in similar ways: the world is expected to contain quantities of things, and we can expect our quantity-based interaction with it to loosely correspond to the output of the ANS, so that, for example, if we see a large quantity of food available somewhere, we can expect to eat a lot by going there.

This explains why infants and animals look longer at impossible outcomes in violation- of-expectancy studies: they have built-in expectations which embody rules about how the world is supposed to behave, and when researchers artificially break these rules, the subjects are struck by the fact that their expectation of how the world works aren't met. The bottom line is that, for both expressed and unexpressed rules, the same feeling of 'fit', or 'ought', guides our behaviour.[13] The difference is that, for expressed rules, we are equipped with a different type of expectation of how the world works, since our expectations of how the world works are built from experience and learning, while for unexpressed rules, our expectations of how the world works is inherited from biological evolution.

If this makes sense, then we can expect to ground explicit rule-following like addition in non-explicit rule-following like allowing oneself to be guided by the content of core cognition modules. Initially, the only rules we expect the world to follow are those we inherit from systems like the ANS. But as we learn more things about the world, we form explicit rules to describe how the world is supposed to behave based on our experience. But the same sense of ought remains, for both unexpressed and expressed rules, based on expectations we have about the world. The upshot would then be that explicit rules depend on unexpressed rules for their existence, but also for their normative character: it is only because we expect the world to behave a certain way that there is a sense of ought attached to rules like addition. Having learned the rule of addition by building on experience and the content of core systems like the ANS, we expect 57 + 68 to make 125, so when someone says that 57 + 68 makes 5, we feel a rule has been broken. This feeling that something isn't as it should be is arguably what we see expressed by animals and infants in violation-of-expectation studies. The difference is, they don't know a rule has been broken.

REFERENCES

Bloor, D. (1997). *Wittgenstein: Rules and Institutions*. London: Routledge. *Davies, P. S. (2000). The nature of natural norms: why selected functions are systemic capacity functions. Noûs 34: 85–107*;

Davies, P. S. (2001). *Norms of Nature: Naturalism and the Nature of Functions*. Cambridge, MA: MIT Press;

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Cohen Kadosh, R. & Dowker, A. (Eds.) (2015) *The Oxford Handbook of Numerical Cognition*. Oxford: Oxford University Press;

Cohen Kadosh, R. & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences*, 32(3–4), 313–28. doi: 10.1017/S0140525X09990938 PMID: 19712504;

13 Hannah Ginsborg's work on primitive normativity (e.g. Ginsborg 2011) seems related to this idea.

De Cruz, Helen. (2008) An Extended Mind Perspective on Natural Number Representation. *Philosophical Psychology* 21, no. 4 : 475–90;
Dehaene, S. & Cohen, L. (2007).Cultural recycling of Cortical maps. *Neuron*, 56 (2), 384-398. 10.1016/j.neuron.207.10.04;
Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press;
Emlen, S. (1975). The stellar orientation system of a migratory bird. *Scientific American,* 233: 102–111;
Feigenson, L., Dehaene, S., and Spelke, E. (2004) Core systems of number. *Trends in Cognitive Sciences* 8:307–14;
Ginsborg, H. (2011). Primitive normativity and skepticism about rules. *Journal of Philosophy* 108 (5):227-254;
Hutto, D. D. & Satne, G. L. (2015). The natural origins of content. *Philosophia (United States)*, 43 (3), 521-536;
Kripke, S. A. (1982). *Wittgenstein on Rules and Private Language*, Cambridge, MA: Harvard University Press;
Kusch, M. (2006). *A sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*, Montréal & Kingston: McGill-Queen's University Press;
Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From 'sense of number' to 'sense of magnitude' - The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*. 1-62. doi:10.1017/S0140525X16000960;
Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 320–334;
Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism.* Cambridge: MIT Press;
Millikan, R. G. (1989a). In defense of proper functions. *Philosophy of Science* 56: 288– 302;
Millikan, R. G. (1989b). Biosemantics. *The Journal of Philosophy* 86 (6): 281–297;
Millikan, R. G. (1990). Truth-rules, hoverflies, and the Kripke-Wittgenstein paradox. *Philosophical Review*, 94 (3): 323-353;
Millikan, R. G. (2005). *Language: A Biological Model*. New York: Oxford University Press;
Odic, D. (2014). *Objects and substances in vision, language, and development* (Doctoral dissertation). Retrieved from https://jscholarship.library.jhu.edu;
Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution.* Chicago: University of Chicago Press;
Rips, L. J. (2017). Core cognition and its aftermath. *Philosophical Topics, 45,* 157- 179;
Shea, N. (2006). Millikan's Contribution to Materialist Philosophy of Mind. *Matière* Première 1: 127-156;
Stern, D. J. (2004). *Wittgenstein's Philosophical Investigations: An Introduction.* New York: Cambridge University Press;
Wittgenstein, L. (2001). *Philosophical Investigations*. Trans G. E. Anscombe. Oxford: Blackwell;
Whiting, D. (2007). The normativity of meaning defended*. Analysis* 67: 133-140;
Strickland, B. (2016). Language reflects "core" cognition: A new theory about the origin of cross-linguistic regularities. *Cognitive Science*. doi: 10.1111/cogs.12332;
Whiting, D. (2013). What is the normativity of meaning? *Inquiry* 1-20.