

## **The Social Origins of Consciousness**

Kristin Andrews<sup>1</sup>, Noam Miller<sup>2</sup>

<sup>1</sup>York University

<sup>2</sup>Wilfrid Laurier University

This is a preprint of an article to appear in *Philosophical Transactions of the Royal Society B*. A link to the final version will be added here when available.

DOI: 10.1098/rstb.2024.0300

## **Abstract**

We present the social origins of consciousness hypothesis, according to which the ability to coordinate with group members was the original adaptive function of consciousness. We offer three arguments. The phylogenetic argument presumes that consciousness is widespread among existing animals, and that widespread capacities are likely evolutionarily old. Early animals relied on consciousness to solve a problem which arose during the Cambrian when animals first became behaviorally flexible-- how to predict others' behavior and stay together as a group. The argument from neuroscience points to evidence that even very simple brains have the capacities for social rewards and pains, and that modern brains retain close connections between the substrates for social cognition and affect. The deep adaptive alignment between social pain and harm to animals develops an argument originally proposed by William James. We provide evidence that in preference tests, bodily pain is preferred to social pain in a wide range of species. We offer two approaches to testing the hypothesis—the salience of social stimuli test and the overattribution of agency test. Working under the social origins of consciousness hypothesis could lead to significant breakthroughs in research, especially by focusing on simpler systems than are currently studied.

**Keywords:** consciousness, sentience, sociality, evolutionary function

## **The social origins of consciousness**

### **1. Why is anything conscious?**

Consciousness is identified first and foremost through first-person experience, not via some externally observable behavior with a clear functional contribution. We feel its existence, but don't know what it's for. Despite this uncertainty, scientific and philosophical consciousness research always involves at least some implicit assumption about its function, and about the observable effects of conscious experience. These assumptions are part of the framework of consciousness research programs.

A familiar presumptive function for consciousness is that it helps individuals avoid painful stimuli and seek out pleasurable ones. Such views tend to emphasize the pain caused by bodily damage, and the pleasures associated with bodily health. Consider William James' adaptive alignment argument against consciousness epiphenomenalism. James observed that what the body needs tends to feel good, and what is harmful to the body tends to feel bad:

“It is a well-known fact that pleasures are generally associated with beneficial, pains with detrimental, experiences... An animal that should take pleasure in a feeling of suffocation would, if that pleasure were efficacious enough to make him immerse his head in water, enjoy a longevity of four or five minutes... if pleasures and pain have no efficacy, one does not see... why the most noxious acts, such as burning, might not give thrills of delight, and the most necessary ones, such as breathing, cause agony” (James 1890/1981, 146).

This focus on bodily health and harms naturally leads to certain research programs and experiments; for example, the current focus on pain in response to tissue damage, shock, or burning acid found in the animal literature reflects this presumption about the function of consciousness. However, this approach may be missing another more primitive function for consciousness, one that focuses on the psychological pains and pleasures associated with sociality.

Here we propose the Social Origins of Consciousness hypothesis. Contemporary science offers us evidence that consciousness may exist widely across animal species (Andrews et al. 2024) and philosophical argument supports the idea that all existing animals are conscious (Andrews 2024). When a capacity is widely distributed, it is likely evolutionarily old. Thus, to ask whether there is some content that can be identified with the original function of consciousness, we turn to examine the point at which consciousness may have first emerged in early evolution. Our proposal is that when behavioral flexibility emerges we find the first instances of true sociality, and that consciousness serves the social function of identifying and responding to the behavior of social partners. This point in evolutionary history is pivotal, because it is the instant at which simpler methods of coordinating behavior no longer suffice.

We offer three arguments for the social origins of consciousness. In §2 we describe the hypothesis, and in §3 we present the first argument, where we develop a phylogenetic story of the origins of consciousness. We ground this phylogeny in known behavioral and anatomical traits of extant taxa, where possible, keeping in mind that such extrapolations can be tenuous. The story begins during the Cambrian explosion (~530 mya), but we first set the stage by describing pre-Cambrian conditions. In §4 we elaborate on this story, suggesting some of the ways in which early brains could have served as substrates for consciousness and examine whether remnants of this early structure can be found in contemporary brains. In §5 we build on James' adaptive alignment argument, reviewing contemporary evidence that existing animals demonstrate a deep affective need for social partners, choosing social contact over basic physical needs. In §6 we propose how this hypothesis could be tested in a wide range of animal species.

Before we begin, a few preliminary remarks. First, we take consciousness to be an umbrella term that encompasses many dimensions, including sentience (feeling pain, pleasure, sensations), imaginings, dreamings, mental time-travel, self-awareness, inner speech, and so forth. With the emergence of any one of these dimensions, a system becomes conscious. In what follows we work from the premise that the earliest type of consciousness was sentience - the capacity for feeling negatively and positively valenced sensations, such as pain and pleasure. We anticipate that the different dimensions of consciousness will have their own adaptive functions, especially when it comes to inner speech, or self-awareness, and that these emerged much later in evolutionary history and are not apparent in all extant species.

Second, our proposal does not address the function of these other dimensions, and does not deny functional pluralism, but is focused on the adaptive advantages associated with the emergence of sentience. This focus shouldn't be taken to imply that sociality is the *only* function of consciousness experience. Consciousness today has a number of different adaptive functions in different domains, which is to be expected given that consciousness today has a large number of different dimensions (Birch et al. 2020; Dung & Newen 2023). This functional pluralism (Ludwig 2023) offers consciousness researchers many paths for investigating consciousness in different domains such as visual processing, emotional expression, memory, and social cognition. We also don't claim that sociality is necessary for consciousness, as the phylogenetic story of life on this planet may not be a universal one. Neither do we claim that an extant species that was truly asocial would lose consciousness, as consciousness would have been exapted to serve other functions after its appearance—again, today's functional pluralism identifies many roles for consciousness to play.

Finally, we note that our proposal takes a very different approach from theories of consciousness which identify some particular neurophysiological structure or process as necessary for consciousness. We take all such accounts to be premature, given that we cannot know what features are necessary for consciousness before knowing which beings are conscious. A function first approach that is open to the possibility that the functional

contributions of consciousness are multiply realizable offers a different path forward for the study of consciousness, and this is the path we take in this paper.

## **2. The Social Origins of Consciousness Hypothesis**

According to the social origins of consciousness hypothesis, the ability to coordinate with group members was the original adaptive benefit for consciousness. Organisms experienced negative feelings when distant from or out of sync with group members, modifying their behavior to regain proximity, and experienced positive feelings when close to the group or coordinating with it. By taking the ability to predict and coordinate behavior with other cognitive beings as a significant step in the evolution of consciousness, we offer a new lens through which to study consciousness in extant organisms.

We intend our story to be an invitation to think more broadly about consciousness as primarily a capacity of social systems (Barlow, 1987). This may seem somewhat paradoxical, since the largest barrier to the scientific study of consciousness comes from its nature as a subjective, first-person phenomenon which is not currently measurable or observable—and hence not socially available. The subjective nature of consciousness may make individualistic approaches to the function of consciousness tempting, but we think this temptation must be resisted. Thus, our proposal is in tension with Cartesian views that take knowledge of the self to be primary, and psychological views that propose sophisticated cognitive capacities such as theory of mind or shared intentionality are needed for true sociality. That is, a familiar story is that the original function of consciousness was to support individual learning about environmental contingencies and that self-consciousness preceded an understanding of other agents as having motivations for their complex behaviors. The social origins of consciousness hypothesis turns this standard view on its head, proposing that consciousness first served to improve predictions about the behaviors of others and only later (possibly much later) was turned inwards toward the self. If anything, the hypothesis suggests that individual action and self-consciousness, or separating oneself from the group, may have been the cognitive achievement.

This story relies on there being a major transition in the dynamics of group living with the emergence of cognitively flexible behavior. This transition involves a separation between the processes of sensation and action, allowing for arbitrary intervals between the two. This gap is filled by the processes we call cognition. Before animals became cognitive beings, their motion was directly and immediately caused by external stimuli, without intervening representations. There was no learning, no memory, and, we are hypothesizing, no consciousness (or at least no adaptive function for consciousness). With the emergence of cognition, individuals' responses to stimuli become more variable and therefore harder for others to predict. In parallel, muscles evolved and replaced cilia as the primary effectors, greatly increasing the scale and speed of action, along with the range of actions possible, which also made prediction more difficult. The emergence of cognition and faster forms of movement thus creates new problems for socially

living organisms, making prior methods for coordinating and predicting others' movements no longer effective. Coordinated behaviors that previously emerged reliably from the aggregative effects of simple slow movement were thrown into disarray by individual variability in history and response biases.

We describe this as a problem for the organisms given that individuals risked losing the benefits of group living, and none of their old solutions sufficed. Consciousness had the function of solving this new problem. However, for it to appear as a problem, the organisms themselves had to care about social partners before they could learn new ways of coordinating with them. Having feelings about one's social situation led animals to focus their attention on social stimuli, and devote a large part of their cognition to social coordination.

### **3. Phylogenetic argument: Sociality and the Cambrian explosion**

The evolutionary history of the Cambrian Explosion (CE) has been interpreted in many different ways, but the general story goes something like this. Along with many other physical and physiological innovations, muscles first emerged at this time, allowing for faster and more vigorous movement, including of larger organisms (Hsieh et al., 2022; van Duijn, 2017). However, muscle movements must be coordinated in order to move the whole organism. This, on many accounts, was the impetus for the evolution of modern neurons (Moroz et al., 2021; though simpler forms predate the Cambrian; Bosch et al., 2017), which coordinated muscles across larger bodies (Keijzer, 2015). Actions also need to be appropriate to incoming sensory information, so neurons also serve to integrate sensations and connect them to action. At about the same time, distal senses - such as image-forming eyes - emerged (Parker, 2003). The combination of distal senses and faster movement enabled predation (Porter, 2011), which in turn led to the evolution of defensive mechanisms (shells and burrowing both emerge in the CE) as well as weapons (hard mouth parts appear in arthropods in the CE, as do the first grabbing appendages; Trestman, 2013). This started the arms race between prey and predators, which partly provided the selective pressure for animals to get larger and better at moving. There are more ways to be successful as a predator, leading to an explosive radiation in body forms and phyla (van Duijn, 2017). It is widely agreed that a corresponding change in behavior and cognition went along with these physiological developments. Importantly for our story, the combination of faster movement at a larger scale with increased processing capacity gave rise to the first behaviorally flexible organisms, heralding an era of extreme unpredictability.

#### *Step 1: before there was consciousness, there was group-living*

Even before the start of the Cambrian era, 650 million years ago, a wide range of Ediacaran animals lived together in groups (e.g., Droser & Gehling, 2008). These animals, with their fractal structures and early experiments with muscles, were attached to the sea bed, as is apparent at

sites such as Mistaken Point Newfoundland, where volcanic ash preserved the ecosystem at a moment in time (Darroch et al., 2013). Thus, group living is ancient. These observations are supported by theoretical arguments that suggest group living was in place early in evolutionary history.

First, consider that today group living is ubiquitous, apparent in every taxon, including bacteria, which create public goods and form cooperative groups (Smith & Schuster, 2019; even pre-Cambrian prokaryotes collectively formed stromatolites, Allwood et al., 2006), single-celled eukaryotes (such as *Dictyostelium*, which displays altruism; Li & Purugganan, 2010; Strassman et al., 2000), plants which communicate chemically about environmental conditions (Wang et al., 2020), and a vast array of animal species (Sumpter, 2006; Krause & Ruxton, 2002). Even animals deemed solitary are coming to be seen as having a non-gregarious form of sociality, or being only opportunistically gregarious, such as the octopuses of Jervis bay (Godfrey-Smith & Lawrence 2012). Many of these groups diverged long before the onset of the Cambrian era. When a phenotype is as widespread and conserved as this one, it is reasonable to hypothesize that it emerged early in evolutionary history.

Second, most early animals – such as ediacarans – had limited (or no) mobility, entailing that individuals remained close to their point of origin, resulting in aggregations of closely related individuals and enhancing opportunities for indirect fitness benefits of cooperation, as observed in bacteria (Belcher et al., 2022). Animals that move passively on ocean currents might also be swept into aggregates, as occurs in extant cnidarians (Graham et al., 2001).

Finally, group living offers benefits to organisms, many of them via mechanisms that require only extremely simple forms of social interaction. For example, predator dilution effects require only that individuals remain close to each other (Hamilton, 1971), as do some public goods in bacteria (Smith & Schuster, 2019). Trilobites aggregated, possibly via shared attraction to chemical cues, aligning the same way relative to a current, or by maintaining physical contact while moving (Vannier et al., 2019), perhaps similarly to how ants (their descendants) employ pheromone trails and tandem running (Möglich et al., 1974). Initial benefits of this sort served to keep animals close together, later allowing for the scaffolding of more complex social interactions and their attendant benefits (“many-eyes” effects, collective predation or defence, predator confusion, etc.; Krause & Ruxton, 2002).

However, before the adaptive radiation of the Cambrian explosion, animals in these groups displayed only simple, inflexible bodily movements. Most importantly, the timescale of sensing and action was slow, as it is in extant adult sponges (Leys et al., 2019). Though animals of this period could likely sense and react to environmental changes and possibly control their own movement to some degree (as in sponges today; Lavrov & Kosevich, 2018), their movement was slow, and their behaviors consisted of direct responses to their environment.

*Step 2: The rise of social uncertainty*

The new more complex mechanisms that evolved early in the Cambrian era - involving more sophisticated sensations collected and processed by increasingly centralized neuronal clusters which then drive coordinated action - along with an increase in the range of possible behaviors, created selective pressures for more processing of information – for cognition. For example, motile predators must be able to subtract their own motion from incoming sensory information to accurately track prey or conspecifics, via a process known as reafference, which may have been one of the earliest functions of neuronal systems (Jékely et al., 2021; Trestman, 2013; Klein & Barron, 2016). The interposition of neurons between sensory and motor systems, and a corresponding temporal gap between sensation and action, marks the origin of cognition.

With the emergence of cognitive processes, organisms began making decisions and incorporating more information into decision-making, including elements retrieved from early forms of memory. Behavior was no longer predictable or transparent, and coordination became a new challenge. Cognitive processing also allows for responses to be delayed relative to their motivating stimuli, which makes prediction even more difficult. With movement no longer transparently and directly caused by environmental factors, organisms no longer had a simple mechanistic solution to coordinate their behaviors. The rules by which organisms moved could vary from individual to individual and be modified over the course of one's lifetime.

At this stage, the very possibility of groups was at risk. If each organism failed to attend to the behavior of their group members and adjust their behavior accordingly, they could have gotten lost in a sea teeming with things to move toward and avoid. Individuals could swim off, move against the current, dart after prey or move away from a predator. What had been slow and predictable became chaotic and unpredictable. Groups would have shattered, along with the adaptive benefits of group living. Just at the point that living in groups would have offered benefits of protection from newly evolved predators, individuals no longer had a method for sticking together. Given the immense value of group living, there was strong selective pressure for something to emerge to support continued collective living.

Cognition and behavioral flexibility created the problem, but it also gave animals the tools they needed to adjust to the increasingly flexible behaviors of others, by learning how to predict more complex behaviors. However, for cognition to work to solve this problem, more of an animal's cognition had to be directed towards the actions of others.

### *Step 3: Closing the predictability gap with consciousness*

At this point, we propose that phenomenal consciousness becomes socially adaptive. On the question of the first spark of sentience, we remain silent. Instead, our focus is on the environmental problem that feelings could uniquely solve.

Those organisms who experienced negative affect when group members were not around were motivated to approach conspecifics. This provided the incentive for animals to focus their attention onto other organisms and their social signals, allowing them to more accurately predict others' behaviors. This process is visible in extant animals, whose attention focusing



skills may develop partly via social mechanisms (e.g., in young humans; Yu & Smith, 2016) and who are hyper-attuned to social cues (e.g., Adolphs & Spezio, 2006; Klein et al., 2009; Rosa Salva et al., 2015).

With its attention focused on its group mates, an organism could detect discrepancies between its predictions of others' actions and their actual behaviors. These gave rise to negative affect, as mismatched behaviors risked dissolving the group, which motivated error correction mechanisms to keep the group together. Animals may also have gradually learned to better predict the behaviors of others in their group, improving future predictions, though we make no claim about when learning emerged. We take this point as the origin of sociality, as distinct from mere group living. Sociality is how behaviorally flexible organisms design their social interactions, be it a gregarious structure in unique species groups, communities of a variety of different organisms as in multi-species bird flocks, or remote and occasional interactions between same-species individuals and regular interaction with other species (Kappeler, 2019). Sociality emerges from other-oriented cognition.

An alternative way to maintain group cohesion and its attendant anti-predatory benefits is to behave in a way that is more easily predictable by other members of the group, though ideally still *unpredictable* to potential predators. The process we describe above, of increasing attention to and processing of social cues, would also have acted to increase the complexity of behaviors animals could engage in without disrupting collective movement. In extant animals, emotional states are often broadcast as signals to conspecifics (e.g., blushing; Leary & Toner, 2012), further enhancing predictability (knowing another's emotional state is key to predicting their future actions). This could be considered an evolutionary version of the mindshaping argument, according to which simple behavioral coordination leads to shared internal states that facilitate more complex coordination (Zawidzki, 2018).

Behavior prediction is an essential skill for cognitive organisms, and the capacities involved in behavior prediction come in a great variety (Andrews 2012). What all the methods of behavior prediction require is some cue which might be taken as a cause of the detected behavior. Importantly, the ability to use affect to learn and to perform error-correction behaviors requires that animals can identify aspects of the environment relative to which the relevant behavior must be organized, and so requires both sensory awareness and valenced affect.

There are three possible stories about the initial appearance of sensory awareness and affective sentience in evolutionary history. It is possible that awareness of sensory stimuli—conspecifics, potential prey and predators, elements of the external world such as light or temperature—evolved first, and only later did organisms gain an ability to attach valenced affect to those conscious representations. Alternatively, animals may have first evolved the ability to have valenced sensations about stimuli that they were not consciously aware of, experiencing a kind of free-floating pleasure or pain, and later developed the ability to be aware of sensations. Finally, the two could also have evolved concurrently.

Though our theory is focused on the emergence of sentience and its adaptive benefits, not sensory awareness, our story is consistent with all three of these possibilities. In any of these scenarios, once behavioral flexibility arose, the old methods of coordinating behaviors no longer sufficed. A new problem emerged, but with it came new ways of solving problems. Organisms could learn new means of coordinating behavior, associating social cues with future actions, and find new ways of keeping groups together. First, however, the organisms needed to value sociality, and recognize the chaotic social world as a problem. We hypothesize that consciousness originally served the function of directing attention to social partners, making sociality feel good, and isolation feel bad.

#### **4. The argument from neuroscience**

If some forms of consciousness, such as affect or sensory awareness, evolved as early as we are proposing, then they are a feature of relatively simple neural systems, which is all that existed in the early Cambrian. Debates around the neural correlates of consciousness mostly concern more complex forms than we are considering. However, it is conceivable that sentience could be supported by very simple brains.

At its core, sensory awareness involves reprocessing of some sensory information (or re-presenting it to the brain; Humphrey, 1987). In other words, awareness is a function of feedback. In addition to processed sensory information being used to drive action-selection, stimuli that we are aware of are reflected back into the brain. Affect provides an evaluative signal for how to respond to the sensory information. It has been suggested that early neural systems had just such a feedback loop, which was used for reafference (Klein & Barron, 2016; reafference is also present in extant sponges and ctenophores, supporting its early origins; Jékely et al., 2021). In reafference, motor information is reprocessed as sensory data to allow for the subtraction of one's own movement from the external sensory stream. It is possible that the neural structures that underlie this feedback loop were duplicated, perhaps via one of the genome duplication events that occurred during the CE (Holland, 2015), and the extra copy was then used as an initial substrate for sensory awareness. In other words, brains did not evolve sentience due to specific selective pressures for consciousness; consciousness resulted from an exaptation of existing brain structures, retained because of the social benefits it made possible.

Modern brains retain close connections between the substrates for social cognition and affect. For example, a wide range of vertebrate species (Bickart et al., 2014), from mammals (Newman, 1999) to reptiles (Skinner et al., 2024) and fish (Bshary et al., 2014), rely on an interconnected set of brain regions to process social information (the "social behavior network"), which partly overlaps with and is strongly connected to the mesolimbic reward system, a key region for emotion (the two areas together are sometimes called the "social decision-making network"; O'Connell & Hoffman, 2012).

Additionally, the mu-opioid system, closely involved in the regulation of pain, also plays a central role in modulating many social behaviors (reviewed in Johnson & Dunbar, 2016). Pain-processing areas of the brain can be activated not only by stimuli that are directly (or potentially) physically harmful but also by social rejection (Eisenberger et al., 2003) or by observing another in pain (Rütgen et al., 2015), and the intensity of such empathetic pain correlates with social closeness to the partner (reviewed in Gu et al., 2020). Empathetic pain activates the same areas of the brain involved in the affective experience of physical pain (Singer et al., 2004), and can be reduced using painkillers (Mischkowski et al., 2016). The oxytocin system is also closely involved in both increasing prosocial behavior (possibly by increasing the salience of social cues; Shamay-Tsoory & Abu-Akel, 2016) and stress reduction (Love, 2018) and may play a role in social buffering (a reduction of stress resulting from social contact; Crockford et al., 2018).

## **5. Deep adaptive alignment argument**

A third argument supporting the social origins of consciousness hypothesis comes from a development of James' adaptive alignment argument, which he gave in response to T.H. Huxley's claim that while bodily changes impact conscious experience, conscious experience never causes changes to bodily states—just as, in his famous metaphor, the steam engine causes the whistle to produce sound, but the sound has no impact on the function of the engine (Huxley 1874/1894). As Huxley was a fellow evolutionary theorist, James' arguments appealed to the shared premise that if consciousness made no causal contribution, it couldn't have evolved through natural selection. James offers three reasons to think that “life-essential, phenomenal pleasures and pains” (Klein 2019) are adaptive: they are complexly organized; they are linked with beneficial and harmful brain and bodily states; and these patterns are universal across humans. There is no other explanation for these three features, thus conscious experiences of pain and pleasure must be causally efficacious, and have been selected for.

While James offers this as an argument against epiphenomenalism, it is also an argument for a function of consciousness. However, James was focused on causes of pain and pleasure in terms of physical damage to the body—for the sake of convenience we will refer to these as “bodily pains”. This focus on bodily pain is made explicit in a laundry list of pain examples given by Grant Allen, a contemporary of James, which includes: pain due to passing gallstones, breaking nails below the quick, having chapped lips, abscesses, ulcers, cancers, eating excess cayenne pepper, destroying the skin and muscles with acid, among many other more or less gruesome experiences (as discussed in Klein 2019, p. 1185). But in this list we do not find the death of one's young child, losing an aged parent who wanders off in the woods, being forsaken by a lover, having a friend commit suicide, being a survivor of an attack that killed your family, or any other kind of social pain. That is, what we can call “social pains” are not considered.

James argues that bodily pains are adaptive not just because of a general adaptive alignment, but also by their deep adaptive alignment, that the more serious the damage, the worse the pain. Passing gallstones is more painful than chapped lips, and gallstones are also more life threatening than chapped lips. Of course, the mapping isn't perfect, and we now enjoy many things that are harmful to us as well, but the idea is that there is a rough mapping of degrees of pain and degrees of harm.

Starting with James' idea of this deep adaptive alignment, we can support the social origins of consciousness hypothesis by showing that social pains are more painful than bodily pains, and that social attention is deeply desired and will be sought out even when it increases the risk of bodily pain. If consciousness was selected for because of its role in facilitating group living at a time of upheaval and increasing unpredictability of individual behavior, we would expect that social ills are the most aversive. Following James' logic, it would be quite a coincidence if most existing animals have deep preferences for social goods, and sociality offers immense evolutionary benefits, and yet consciousness was not an early adaptation for social cohesion.

While it has been well established that social isolation causes significant psychological and physical damage, perhaps the strongest evidence comes from preference tests which support the claim that social attention is deeply preferred across a wide range of species, and social isolation is among the most aversive experiences.

One piece of evidence for the relative importance of social attention comes from studies of the importance of social contact. Among the earliest were Harry Harlow's studies of primates in the 1960's where he deprived infants of maternal care and socially isolated adult individuals. Harlow's goal was to better understand human mental disorders by intentionally creating disorders in monkeys and chimpanzees. Infant monkeys were separated from their mothers for 6 to 12 months and provided a choice between two objects, one a mesh wire "mother" that delivered milk, the other that was soft but delivered no food. The infants preferred the soft "mother", speaking to the deep importance of social interaction and touch in primates. Indeed, when infant monkeys were separated into two conditions, one in which they only had access to the wire "mother" and the other in which they only had access to the cloth "mother", infants in the wire "mother" context showed significant psychological deficits (Harlow 1958). Subsequent research on the importance of touch in mammals have discovered the existence of nerve fibers called C-tactile afferents that are specially attuned to social touch, particularly affiliative touch such as gentle stroking (Löken et al. 2009; McGlone et al. 2014). These have been found in all mammals studied, including primates, pigs, rats, mice, guinea pigs, rabbits, and cats (Morrison 2012; Pitcher et al. 2016). Social touch has been identified as a significant moral interest for all animals (Monsó & Wrangé 2019).

Social enrichment is widely understood to be necessary for health and wellbeing in mammals, and there is a growing awareness of its importance to reptiles and amphibians (Nagabaskaran et al., 2022; Warwick et al., 2023). In formal studies, hamsters will push heavy doors to gain social contact (Borland et al. 2017), rats prefer social contact to chocolate (Bartal et al. 2016),

rhesus monkeys will starve themselves rather than giving a conspecific a shock (Wechkin et al. 1964), and mice prefer a cagemate over nesting material (Van Loo et al. 2004). A wide variety of animal models have been studied to show that sociality lowers self-administration of drugs such as cocaine, ethanol, and opiates (Neiswander et al. 2012), and even fruit flies have been found to prefer healthy foods over alcohol after social sexual interactions (Shohat-Ophir et al. 2012). The interpretation is further supported by studies of trade-offs between social interaction and physical pain. In one study, trout, but not goldfish, were found to prefer social contact to avoiding a shock; the trout will suffer increasingly strong shocks to approach a social partner (Dunlop et al. 2006). Lack of social housing in rats has been shown to impact research results on drug trials, leading to significant problems with translation to humans and failures of replicability (Cait et al. 2024). Of course, more research will be informative, but together these studies provide corroborative evidence supporting the claim that animals have a deep need for social contact and value it highly enough to (temporarily) forgo even basic necessities such as food and shelter.

## **6. Testing the Social Origins of Consciousness Hypothesis**

Though we obviously cannot examine the behaviors and cognition of early Cambrian animals, who we propose were already conscious, the social origins of consciousness hypothesis makes some predictions about the structure of consciousness that, if our story is correct, may be observable in extant species as well. We identify two types of tests which can help to corroborate, or undermine, the social origins of consciousness hypothesis. These tests should be conducted on a wide range of animal species, not just the primates so often used in contemporary consciousness research, but also non-mammalian vertebrates and invertebrate species (Miller 2025).

*The salience of social stimuli test.* If consciousness evolved to support social cohesion, then we should expect that social cues should be disproportionately salient compared with non-social cues. There is some evidence of this in the human literature, for example in the propensity to focus on the eyes of others (e.g., Birmingham et al., 2009) and use their gaze direction to extract social information (Gallup et al., 2012; Birmingham et al., 2008). This suggests that humans are hard-wired to attend specifically to those features of the social environment that are likely to help predict the future actions of others (for example, faces showing an emotional expression attract more attention; Vuilleumier, 2002), and there is some evidence that reactions to such social cues engage different brain regions than non-social spatial cues (Greene et al., 2009). Though this effect is congruent with our story, it has only been tested in a few species, mostly humans and other mammals. According to our story, the origins of this effect are ancient and should be observable in a very wide range of species, including those often considered non-social.

The hypothesis that social cues are inherently better at attracting attention could be tested in several ways. For example, humans (and other animals) should be more attracted to social cues that are predictive of future behavior than those that are not, and this distinction may be learned. Masking effects should be stronger when a social stimulus is used as the mask, or weaker when the target to be masked is a social cue. Similarly, since salience is well-known to affect learning, learning that relies on social cues should proceed faster or be remembered better than similar tasks that involve non-social cues.

*The overattribution of agency test.* Agency detection is taken to be innate for many animals (Carey, 2009; Spelke 2022). Our innate focus on social cues may relate to humans' propensity to anthropomorphize - viewing even non-social things in the environment as if they were social, using our social skills to make predictions about, for example, the movements of objects (such as in the Heider-Simmel animation; e.g., Torabian & Grossman, 2023). This research is often characterized as leading to an error—the tendency to overattribute human characteristics to nonhumans, including faces on buildings and garbage cans as well as humanlike motivations or social relationships to other species. However, if all animals are specially attuned to social stimuli, anthropomorphism may simply be the human version of a more general feature, namely the overattribution of agency.

Darwin wrote about the overattribution of agency in animals when he discussed his dog's reaction to an open parasol which was occasionally moving in the wind (Darwin 1871). Corroborating Darwin's report, in experimental contexts dogs appear to attribute animacy to objects that behave in seemingly agentic ways (Gergely et al., 2013). While there have been a few studies of overattribution of agency in animals, such as a null study of rhesus monkeys watching a Heider/Simmel type animation (Schafroth et al. 2021), there is very little evidence either way that animals across species overattribute agency. Research on agency attribution could be conducted in a wide range of species, with particular stimuli calibrated to the species' ecological niches and sensory systems. Visual stimuli watched on a computer monitor may not be equally salient to all animal species.

In sum, we suggest that a more robust study of sociality in a wide range of animal species—going far beyond the typical monkey subjects of consciousness research—would help to support or undermine the social origins of consciousness hypothesis.

## **7. Conclusion**

In this presentation and defense of the social origins of consciousness hypothesis, we have focused on the adaptive function of sentience. Our suggestion is that social cognition, and more specifically the capacity to coordinate with group members after the rise of cognition, was the original function of consciousness, and as such should be expected to be almost ubiquitous in existing conscious animals. One implication is that self-awareness, rather than understanding of others, should be taken as the major cognitive achievement.

Our three arguments for the hypothesis draw on the observation that sociality is evolutionarily ancient and essential to animals. Threats to sociality require solutions, and one of those solutions will be an emotion-driven motivation to be in the presence of social partners. In modern animals, the neurophysiological structures that regulate pain, a basic type of sentience, also work to process social information and modulate social responses. These physiologically connections are reflected in the widespread preference for physical pain over social pain, and the robust evidence that social attention modulates physical pain.

We hope to inspire testing of the hypothesis, but also to inspire including a wider range of animal subjects in the science of consciousness. If consciousness emerged to promote sociality, the study of consciousness in social animals already widely used in research, such as nematode worms and fruit flies, would be a promising means of making progress on discovering those properties that are necessary for conscious experience. We also offer a new focus in research on consciousness, one that places attention on the types of stimuli used rather than on the sophistication of the subject's sensory modalities or cognitive capacities. Testing consciousness using social stimuli in a wide range of social species, may allow us to uncover its original function and take some tentative steps towards understanding its mechanisms.

## References

- Adolphs R, Spezio M (2006). Role of the amygdala in processing visual social stimuli. *Progress in Brain Research*, 156:363-378.
- Allwood AC, Walter MR, Kamber BS, Marshall CP, Burch IW (2006). Stromatolite reef from the Early Archaean era of Australia. *Nature*, 441:714-718.
- Andrews, K. (2012). *Do Apes Read Minds? Toward a New Folk Psychology*. Cambridge: MIT press.
- Andrews, K. (2024). "All animals are conscious": Shifting the null hypothesis in consciousness science. *Mind & Language*, n/a(n/a). <https://doi.org/10.1111/mila.12498>
- Andrews, K., Birch, J., Sebo, J., and Sims, T. (2024) *Background to the New York Declaration on Animal Consciousness*. nydeclaration.com.
- Barlow H (1987). The biological role of consciousness. In: C. Blakemore & S. Greenfield (Eds.), *Mindwaves*. Oxford: Basil Blackwell.
- Bartal, I. B., Shan, H., Molasky, N. M. R., Murray, T. M., Williams, J. Z., Decety, J., & Mason, P. (2016). Anxiolytic Treatment Impairs Helping Behavior in Rats. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00850>
- Belcher LJ, Dewar AE, Ghouli M, West SA (2022). Kin selection for cooperation in natural bacterial populations. *PNAS*, 119:e2119070119.
- Bickart KC, Dickerson BC, Feldman Barrett L (2014). The amygdala as a hub in brain networks that support social life. *Neuropsychologia* 63:235–248.
- Birch, J., Schnell, A. K., & Clayton, N. S. (2020). Dimensions of animal consciousness. *Trends in Cognitive Sciences*, 24(10), 789–801. <https://doi.org/10.1016/j.tics.2020.07.007>
- Birmingham E, Bischof WF, Kingstone A (2008). Social Attention and Real-World Scenes: The Roles of Action, Competition and Social Content. *Quarterly Journal of Experimental Psychology*, 61:986-998.
- Birmingham E, Bischof WF, Kingstone A (2009). Get real! Resolving the debate about equivalent social stimuli. *Visual Cognition*, 17:904-924.
- Borland JM, Frantz KJ, Aiani LM, Grantham KN, Song Z, Albers HE (2017). A novel operant task to assess social reward and motivation in rodents. *Journal of Neuroscience Methods*, 287:80-88.
- Bosch TCG, Klimovich A, Domazet-Lozo, T, Gruender S, Holstein TW, Jekely G, et al. (2017). Back to the basics: cnidarians start to fire. *Trends in Neurosciences*, 40:92-105.



- Bshary R, Gingins S, Vail AL (2014). Social cognition in fishes. *Trends in Cognitive Sciences*, 18:465-71.
- Cait, J., Avey, M. T., & Mason, G. J. (2024). The impact of social isolation on laboratory rodent health: a protocol for a systematic review. *Applied Animal Behaviour Science*, 278:106361.
- Carey, S. (2009). *The Origin of Concepts* (Reprint edition). Oxford University Press.
- Clark RE, Manns JR, Squire LR (2002). Classical conditioning, awareness, and brain systems. *TICS*, 6:524-531.
- Crockford C, Deschner T, Wittig RM (2018). The role of oxytocin in social buffering: what to primate studies add? *Curr Top Behav Neurosci*, 35:155-173.
- Darroch SAF, Laflamme M, Clapham ME (2013). Population structure of the oldest known macroscopic communities from Mistaken Point, Newfoundland. *Paleobiology*, 39:591-608.
- Darwin, C. (1871). *The Descent of Man*. Dover.
- Droser ML, Gehling JG (2008). Synchronous Aggregate Growth in an Abundant New Ediacaran Tubular Organism. *Science*, 319:1660-1662.
- Dung, L., & Newen, A. (2023). Profiles of animal consciousness: A species-sensitive, two-tier account to quality and distribution. *Cognition*, 235, 105409.  
<https://doi.org/10.1016/j.cognition.2023.105409>
- Dunlop, R., Millsopp, S., & Laming, P. (2006). Avoidance learning in goldfish (*Carassius auratus*) and trout (*Oncorhynchus mykiss*) and implications for pain perception. *Applied Animal Behaviour Science*, 97(2), 255–271. <https://doi.org/10.1016/j.applanim.2005.06.018>
- Dunn CW, Leys SP, Haddock SHD (2015). The hidden biology of sponges and ctenophores. *Trends in Ecology & Evolution*, 30:282-291.
- Eisenberger NI, Lieberman MN, Williams KD (2003). Does Rejection Hurt? An fMRI Study of Social Exclusion. *Science*, 302:290-292.
- Gallup AC, Hale JJ, Sumpter DJT, Couzin ID (2012). Visual attention and the acquisition of information in human crowds. *PNAS*, 109:7245-7250.
- Gergely A, Petro E, Topal J, Miklosi A (2013). What Are You or Who Are You? The Emergence of Social Interaction between Dog and an Unidentified Moving Object (UMO). *PLoS One*, 8:72727.
- Godfrey-Smith, P., & Lawrence, M. (n.d.). Long-term high-density occupation of a site by Octopus tetricus and possible site modification due to foraging behavior. *Marine and Freshwater Behaviour and Physiology*, 45(4), 1–8.  
<https://doi.org/10.1080/10236244.2012.727617>

Graham WM, Pages F, Hamner WM (2001). A physical context for gelatinous zooplankton aggregations: a review. *Hydrobiologia*, 451: 199-212.

Greene DJ, Mooshagian E, Kaplan JT, Zaidel E, Iacoboni M (2009). The neural correlates of social attention: automatic orienting to social and nonsocial cues. *Psychological Research PRPF*, 73:499-511.

Gu R, Liu J, Cui F (2019). Pain and social decision-making: New insights from the social framing effect. *Brain Science Advances*, 5:221-238.

Hamilton WD (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, 31:295-311.

Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13(12), 673–685.  
<https://doi.org/10.1037/h0047884>

Holland PWH (2015). Did homeobox gene duplications contribute to the Cambrian explosion? *Zoological Letters*, 1:1.

Hsieh S, Plotnik RE, Bush AM (2022). The Phanerozoic aftermath of the Cambrian information revolution: sensory and cognitive complexity in marine faunas. *Paleobiology*, 48:397-419.

Humphrey N (1987). The inner eye of consciousness. In: C. Blakemore & S. Greenfield (Eds.), *Mindwaves*. Oxford: Basil Blackwell.

Huxley, Thomas Henry. 1874/1894. "On the Hypothesis That Animals Are Automata, and Its History." In *Collected Essays: Method and Results*, 199–250. New York: Appleton.

James, W (1981). *The Principles of Psychology*. Harvard University Press.

Johnson KV-A, Dunbar RIM (2016). Pain tolerance predicts human social network size. *Scientific Reports*, 6:25267.

Jékely G, Godfrey-Smith P, Kieijzer F (2021). Reafference and the origin of the self in early nervous system evolution. *Phil. Trans. B*, 376:20190764.

Kappeler PM (2019). A framework for studying social complexity. *Behavioral Ecology & Sociobiology*, 73:13.

Keijzer F (2015). Moving and sensing without input and output: early nervous systems and the origins of the animal sensorimotor organization. *Biology & Philosophy*, 30:311-331.

Klein, A. (2019). William James's Objection to Epiphenomenalism. *Philosophy of Science*, 86(5), 1179–1190. doi:10.1086/705477

Klein JT, Shepherd SV, Platt ML (2009). Social attention and the brain. *Current Biology*, 19:R958-R962.

Krause J, Ruxton GD (2002). *Living in Groups*. Oxford University Press.

Lavrov AI, Kosevich IA (2018). Stolonial Movement: A New Type of Whole-Organism Behavior in Porifera. *Biological Bulletin*, 234:58–67.

Lays SP, Mah JL, McGill PR, Hamonic L, De Leo FC, Kahn AS (2019). Sponge Behavior and the Chemical Basis of Responses: A Post-Genomic View. *Integrative and Comparative Biology*, 59:751–764.

Leary MR, Toner K (2012). Psychological theories of blushing. In WR Crozier, PJ de Jong (Eds.), *The Psychological Significance of the Blush*. Cambridge University Press.

Li SI, Purugganan MD (2010). The cooperative amoeba: *Dictyostelium* as a model for social evolution. *Trends in Genetics*, 27:48-54.

Löken, L., Wessberg, J., Morrison, I., McGlone, F., & Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nature Neuroscience*, 12(5), 547–548.  
10.1038/nn.2312

Love TM (2018). The impact of oxytocin on stress: the role of sex. *Curr Opin Behav Sci*, 23:136-142.

Ludwig D (2022). The functional contributions of consciousness. *Conscious and Cognition*, 104:103383. doi: 10.1016/j.concog.2022.

McGlone, F., Wessberg, J., & Olausson, H. (2014). Discriminative and affective touch: Sensing and feeling. *Neuron*, 82(4), 737–755. 10.1016/j.neuron.2014.05.001

Miller, N. 2025. Why we should study animal consciousness. *Learn Behav*, 53:3-4.

Mischkowski D, Crocker J, Way BM (2016). From painkiller to empathy killer: acetaminophen (paracetamol) reduces empathy for pain. *Social cognitive and affective neuroscience*, 11:1345–1353.

Möglich M, Maschwitz U, Hölldobler B (1974). Tandem calling: a new kind of signal in ant communication. *Science*, 186:1046-1047.

Monsó S & Wrage B (2021) Tactful animals: How the study of touch can inform the animal morality debate, *Philosophical Psychology*, 34:1, 1-27, DOI: 10.1080/09515089.2020.1859100

Moroz LL, Romanova DY, Kohn AB (2021). Neural versus alternative integrative systems: molecular insights into origins of neurotransmitters. *Phil. Trans. R. Soc. B*, 376:20190762.

Morrison, I. (2012). CT afferents. *Current Biology*, 22(3), R77–78.

Nagabaskaran G, Skinner M, Miller N (2022). Western Hognose Snakes (*Heterodon nasicus*) Prefer Environmental Enrichment. *Animals*, 12:3347.

Nanay B (2013). *Between perception and action*. Oxford University Press.

Neisewander, J. L., Peartree, N. A., & Pentkowski, N. S. (2012). Emotional valence and context of social influences on drug abuse-related behavior in animal models of social stress and prosocial interaction. *Psychopharmacology*, 224(1), 33. <https://doi.org/10.1007/s00213-012-2853-3>

Newman SW (1999). The Medial Extended Amygdala in Male Reproductive Behavior: A Node in the Mammalian Social Behavior Network. *Ann. N. Y. Acad. Sci.*, 877:242–257.

O’Connell LA, Hoffmann HA (2012). Evolution of a vertebrate social decision-making network. *Science*, 336:1154-1157.

Parker A (2003). *In the Blink of an Eye*. Perseus Publishing, 316 pp.

Pitcher, M., Le Pichon, C., & Chesler, A. (2016). Functional properties of C-Low Threshold Mechanoreceptors (C-LTMRs) in nonhuman mammals. In H. Olausson, J. Wessberg, I. Morrison, & F. McGlone (Eds.), *Affective Touch and the Neurophysiology of CT Afferents* (pp. 31–48). Springer New York.

Porter S (2011). The rise of predators. *Geology*, 39: 607–608.  
<https://doi.org/10.1130/focus062011.1>

Rosa Salva O, Mayer U, Vallortigara G (2015). Roots of a social brain: developmental models of emerging animacy-detection mechanisms. *Neuroscience & Biobehavioral Reviews*, 50:150-168.

Rütgen M, Seidel EM, Silani G, Lamm C (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *PNAS*, 112:E5638-E5646

Schafroth, J.L., Basile, B.M., Martin, A. *et al.* No evidence that monkeys attribute mental states to animated shapes in the Heider–Simmel videos. *Sci Rep* 11, 3050 (2021).  
<https://doi.org/10.1038/s41598-021-82702-6>

Shamay-Tsoory SG, Abu-Akel A (2016). The social salience hypothesis of oxytocin. *Biological Psychiatry*, 79:194-202.

Shohat-Ophir, G., Kaun, K. R., Azanchi, R. & Heberlein, U. *Science* 335, 1351–1355 (2012).

Singer T, Seymour B, O’Doherty J, Kaube H, Dolan RJ, Frith CD (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303:1157-1162.

Skinner M, Daanish D, Damphousse C, Krohmer RW, Mallet P, McKay B, Miller N (2024). Socially-mediated activation in the snake social-decision-making network. *Behavioral Brain Research*, 465:114965.

Smith P, Schuster M (2019). Public goods and cheating in microbes. *Current Biology*, 29:R442-R447.

Spelke, ES. (2022). *What Babies Know: Core Knowledge and Concepts Volume 1*, Oxford: Oxford University Press.

Strassman JE, Zhu Y, Queller DC (2000). Altruism and cheating in the social amoeba *Dictyostelium discoideum*. *Nature*, 408:965-967.

Sumpter DJT (2005). The principles of collective animal behaviour. *Phil Trans R Soc B*, 361:5–22.

Torabian S, Grossman ED (2023). When shapes are more than shapes: perceptual, developmental, and neurophysiological basis for attributions of animacy and theory of mind. *Frontiers in Psychology*, 14:1168739.

Trestman M (2013). The Cambrian explosion and the origins of embodied cognition. *Biological Theory*, 8:80-92.

van Duijn M (2017). Phylogenetic origins of biological cognition: convergent patterns in the early evolution of learning. *Interface Focus*, 7:20160158.

Van Loo PL, Van de Weerd HA, Van Zutphen L, Baumans V. Preference for social contact versus environmental enrichment in male laboratory mice. *Laboratory animals*. 2004;38(2):178-88.

Vannier J, Vidal M, Marchant R, El Hariri K, Kouraiss K, Pittet B, El Albani A, Mazurier A, Martin E (2019). Collective behavior in 480-million-year-old trilobites arthropods from Morocco. *Scientific Reports*, 9:14941.

Vuilleumier P (2002). Facial expression and selective attention. *Current Opinion in Psychiatry*, 15:291-300.

Wang N-Q, Kong C-H, Wang P, Meiners SJ (2020). Root exudate signals in plant-plant interactions. *Plant, Cell & Environment*, 44:1044-1058.

Warwick C, Arena PC, Burghardt GM (2023). (eds.) *Health and care of captive reptiles*. Springer Nature.

Wechkin, S., Masserman, J. H., & Terris Jr., W. (1964). Shock to a conspecific as an aversive stimulus. *Psychonomic Science*, 1(2), 47–48. <https://doi.org/10.3758/BF03342783>

Yu C, Smith LB (2016). The Social Origins of Sustained Attention in One-Year-Old Human Infants. *Current Biology*, 26:1235-1240.

Zawidzki TW (2018). Mindshaping. In: Albert Newen, Leon De Bruin, Shaun Gallagher (eds.) *The Oxford Handbook of 4E Cognition*. Oxford Academic.

**Funding Statement**

Kristin Andrews was supported by the CIFAR Future Flourishing program, York University's York Research Chair program, and SSHRC Insight Grant # 435-2022-0749. Noam Miller was supported by an NSERC Discovery Grant #2023-05765.

**Acknowledgements**

Thanks to helpful conversations with Jake Beck, Tyler Delmore, and Kevin Lande.