

ON THE NATURAL ORIGINS AND FUNCTION OF SELF-CONSCIOUSNESS

LESSONS FROM STUDIES OF ANIMAL TOOL USE

Michael J. Hegarty, PhD



Practically Rational Animals

Michael J. Hegarty

Abstract: This paper defends the conditional thesis: if non-human animals engage in instrumental practical reasoning involving individual innovation, then such animals are likely self-conscious. Against neo-Kantian views that tie rationality to conceptual self-consciousness, and against empiricist critiques that decouple animal rationality from self-consciousness, I argue for an intermediate position. Drawing on cases of animal tool use (hook-bending in birds), I show that representing non-actual states of affairs and combining such representations with goals—a hallmark of individual innovation—requires a form of active cognition best explained by positing nonconceptual self-representation. I distinguish this from teleofunctionalist explanations, which attribute representational combination to selection mechanisms (evolutionary, social, etc.) rather than individual cognition. Where teleofunctional accounts fall short, self-representation plays an explanatory role akin to the essential indexical in human action. The upshot is a novel case for the essential connection between self-consciousness and rationality that accommodates animal cognition while preserving key neo-Kantian insights.

1 Introduction

Some philosophers insist on an essential connection between rationality and self-consciousness (Sellars 1968; McDowell 1996; Moran 2001; Marcus 2012, 2021; Boyle 2012, 2024). One intuition behind this is that a rational thinker has a special kind of active control over what it thinks and does. In neo-Kantian terms, a rational thinker is capable of *spontaneity*: thought and action that is not determined by forces external to the thinker (Moran 2018). For neo-Kantian philosophers like Matthew Boyle and Richard Moran, this spontaneity requires self-awareness because, they say, awareness of oneself as the subject of thinking is required for rational responsiveness to reasons—something possible only for minds capable of spontaneity. But

Res Philosophica, *Vol. 102, No. 3, July 2025, pp. 321–350*
https://doi.org/10.5840/resphilosophica2742
© 2025 Michael J. Hegarty • © 2025 Res Philosophica

Phenomenology and the Cognitive Sciences
https://doi.org/10.1007/s11097-026-10182-y

Cocka-tooling around: innovation and self-representation in avian tool use

Michael J. Hegarty[1]

Accepted: 26 June 2026
© The Author(s), under exclusive licence to Springer Nature B.V. 2026

Abstract

I argue that Goffin's cockatoos are individual innovators: they arrive at novel solutions to some practical problems in thought. Drawing on empirical studies of tool fabrication and use, I defend this claim by ruling out cognitively deflationary alternatives, with particular focus on teleofunctionalist accounts. I then explore what kind of cognition would support such innovation and argue that a minimal form of self-consciousness—nonconceptual self-representation—likely plays a role. By drawing on phenomenological work on the role of the body in cognition—and, especially, in imagination—I suggest that the tool use requires the agent to keep



SUMMARY

- Goffin's cockatoos are capable of a kind of 'individual innovation' in tool use (Hegarty 2026).
- To be capable of this, I argue, it is likely that they are self-conscious in the minimal sense of being capable of *nonconceptual self-representation* (or, perhaps better: *partially conceptual*).
- I remain neutral on the metaphysics of the self. The idea is to **characterise** a representational capacity required to explain a kind of empirically evidenced cognition. I think this capacity qualifies as a kind of self-consciousness because it involves self-representational content in conscious states.
- Self-representational content: the **identity** of body as locus of perception and action **across two egocentric perspectives**—one *tied to present circumstances*, and the other *prospective* (representing merely possible states of affairs).

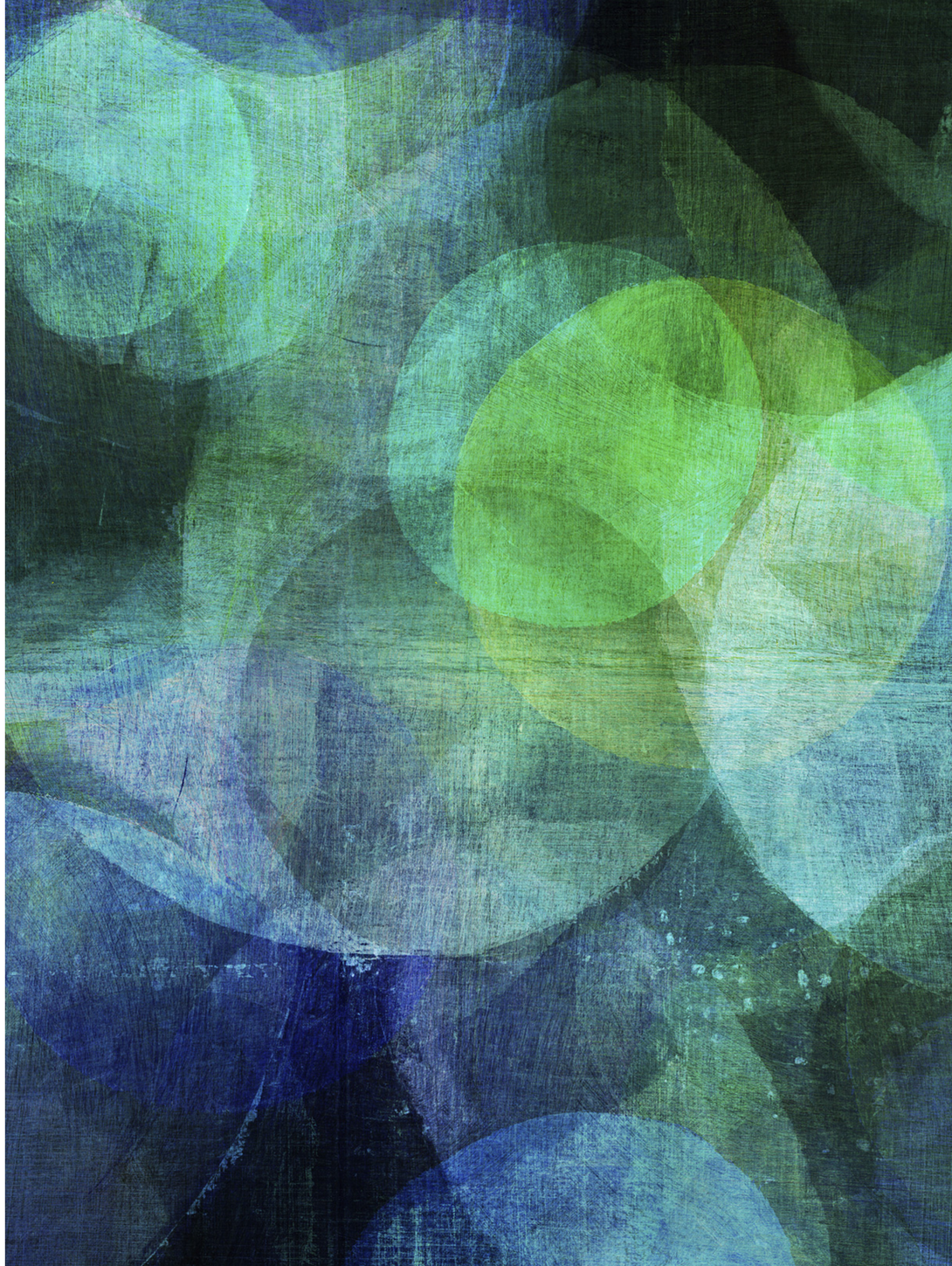


OUTLINE OF QUESTIONS & STRATEGY

- Basic Kantian idea: consciousness always involves(?) self-consciousness (in some sense).
- Roughly: some pre-reflective, ‘self-appearance’—a ‘for-me-ness’ of experience related to the first-person perspective—making possible explicit ‘I’-thoughts (at least in conceptual thinkers), e.g., ‘I am seeing a tree’.
- Can we make this idea more precise?
- Can we defend the Kantian idea by using, in part, empirical materials?
- Animals are plausibly subjects of experience, with points of view. Perhaps they have a kind of pre-reflective self-consciousness?

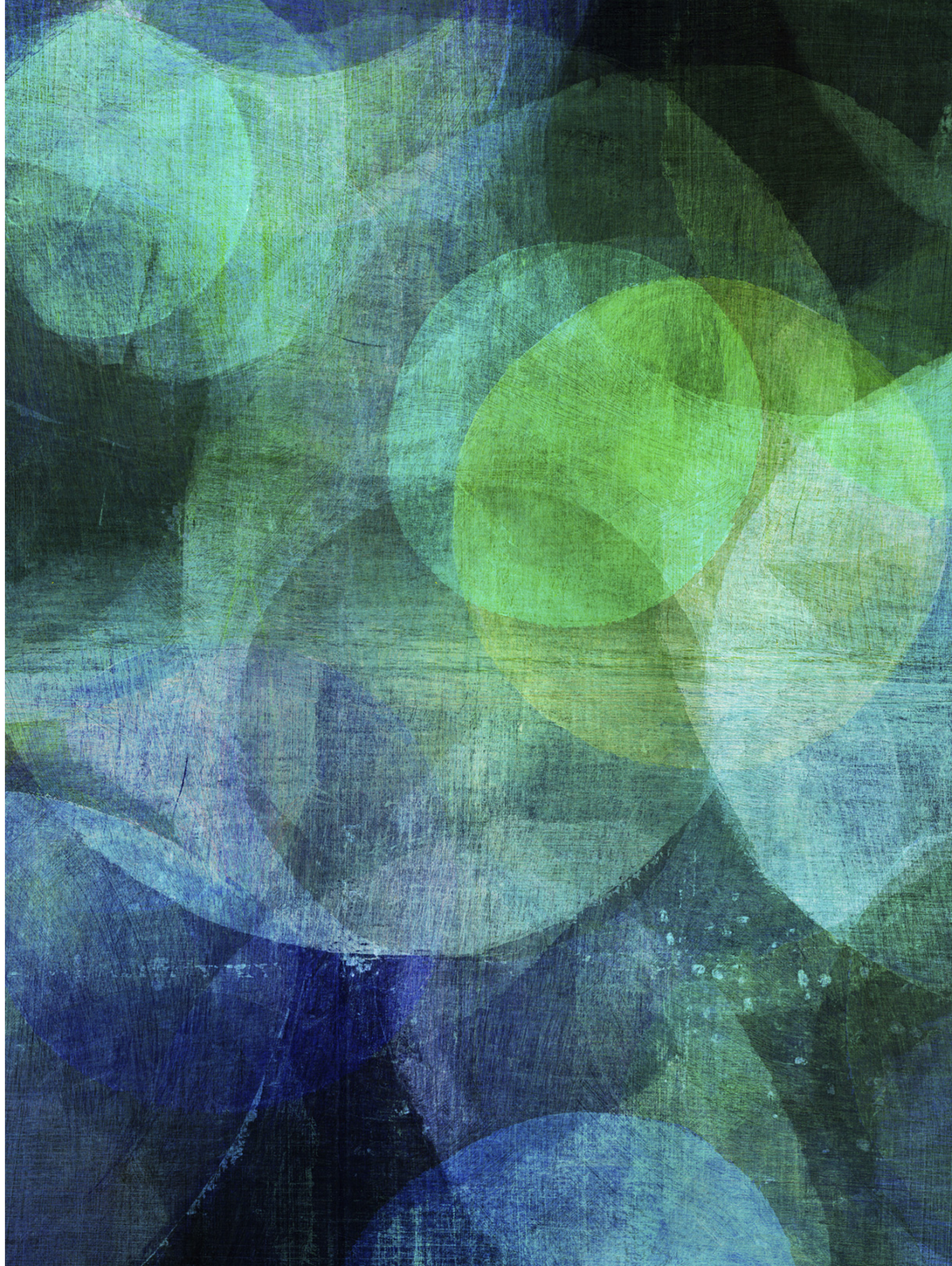
WHY, MORE PRECISELY?

- Many *post-Kantians* understand the Kantian idea as making possible a *unified subject* that is a requirement for *human rational thought*. (E.g., Marcus 2021, Boyle 2025)
- Separately, many *analytic* philosophers think that a first-personal content (*de se* content, the so-called 'essential indexical') common to different thoughts (e.g., intentions and beliefs about ends) is necessary to explain intentional action. (E.g. Perry 1979, 2002; Babb 2016)
- I think we have an analogous situation for the cockatoo: it is engaging in a *limited kind of rationality* (Hegarty 2025) that *combines different thoughts* that *result in intentional action*. So, analogously, we need to explain the unity of subjecthood making that possible.
- Both *post-Kantian* and *analytic* philosophers broadly agree that self-consciousness/self-representation solves their respective problems. So, why not an analogous solution for the cockatoo?



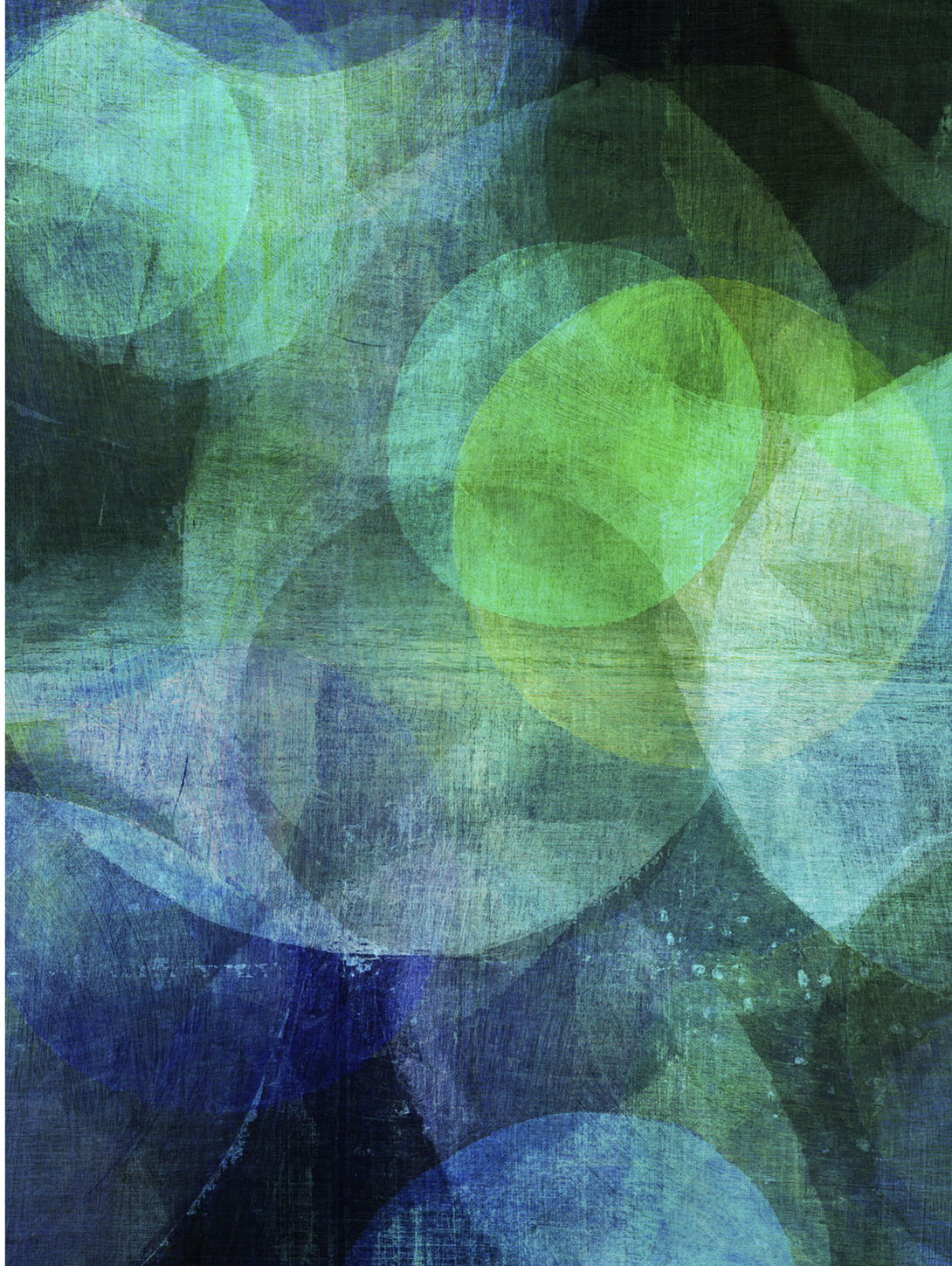
PHILOSOPHICAL BACKGROUND I

- Some follow Kant in thinking "consciousness always involves a self-appearance" (Gallagher & Zahavi, 2025).
- Many *analytic* philosophers think: self-consciousness is the *conceptual capacity to think of oneself as oneself*.
- E.g., 'I am seeing a tree'.
- Self-conscious thinkers then:
    1. Think conceptually.
    2. Have thoughts with propositional (roughly: linguistic) structure.
    3. Possess a SELF concept (perhaps with descriptive content).



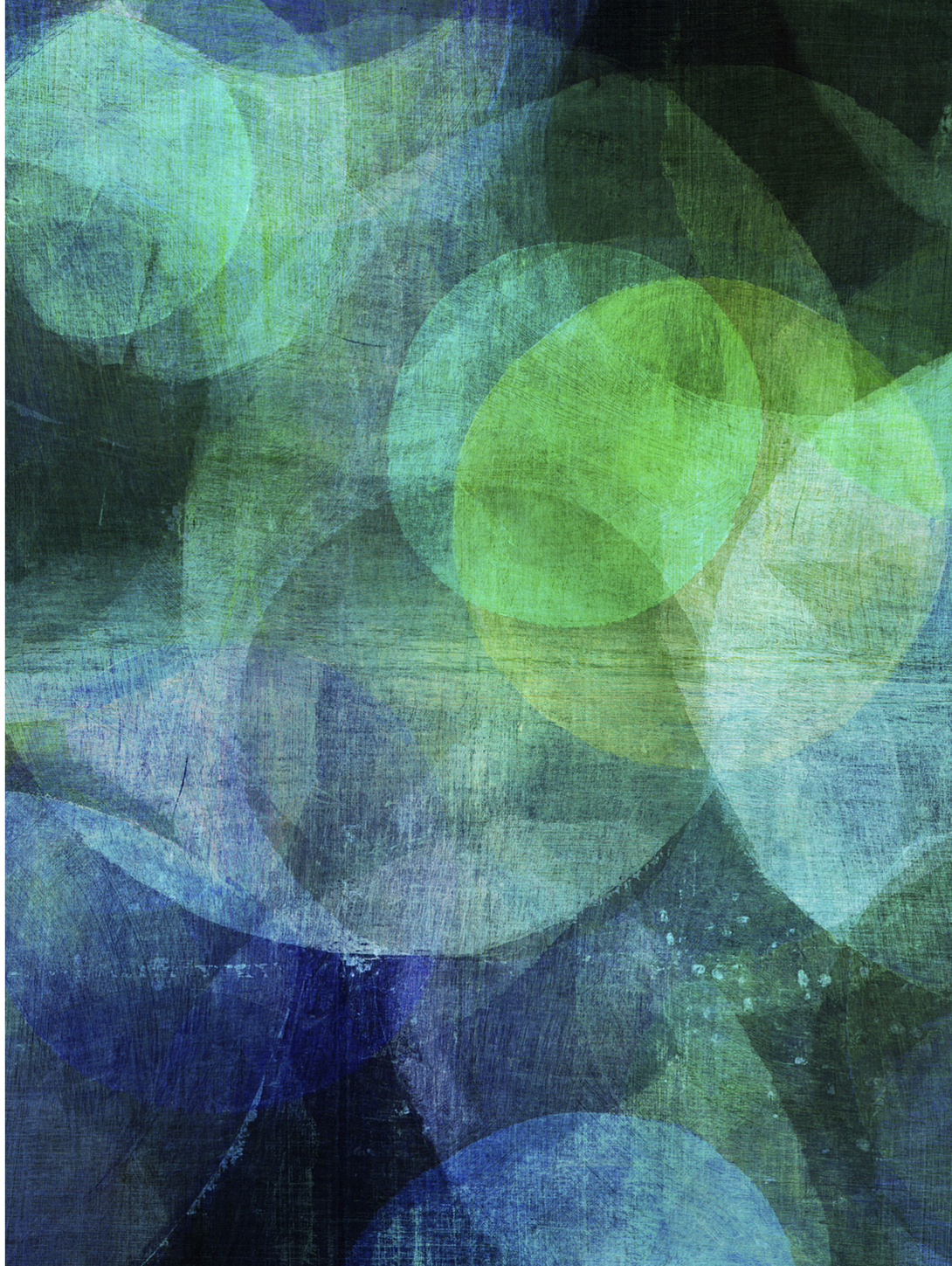
PHILOSOPHICAL BACKGROUND II

- But wait… does consciousness in animals require self-consciousness (in the ‘I’-thoughts sense)?
- Surely a dog, say, is consciously aware of its surroundings, or there are subjective (phenomenal) qualities to its seeing, smelling etc.
- But a dog probably can’t think conceptual-propositional thoughts. It probably has no SELF concept.
- So, the Kantian idea looks dubious.



PHILOSOPHICAL BACKGROUND III

- But *post-Kantian* philosophers—particularly *phenomenologists*—think there is a more basic kind of self-consciousness.
- Perhaps there is a "self-consciousness which is present whenever I am living through or undergoing an experience, e.g., whenever I am consciously perceiving the world, remembering a past event, imagining a future event, thinking an occurrent thought, or feeling sad or happy, thirsty or in pain" (Gallagher & Zahavi, 2025).
- 'Pre-reflective self-consciousness': a kind of 'for-me-ness' of experience that makes possible the self-ascription of states and experiences in 'I'-thoughts such as 'I am seeing a tree'.
- Connected to one's awareness of being (in some sense) a subject of experience, of having a point of view.
- I find this frustratingly vague, yet I think there's something to it.

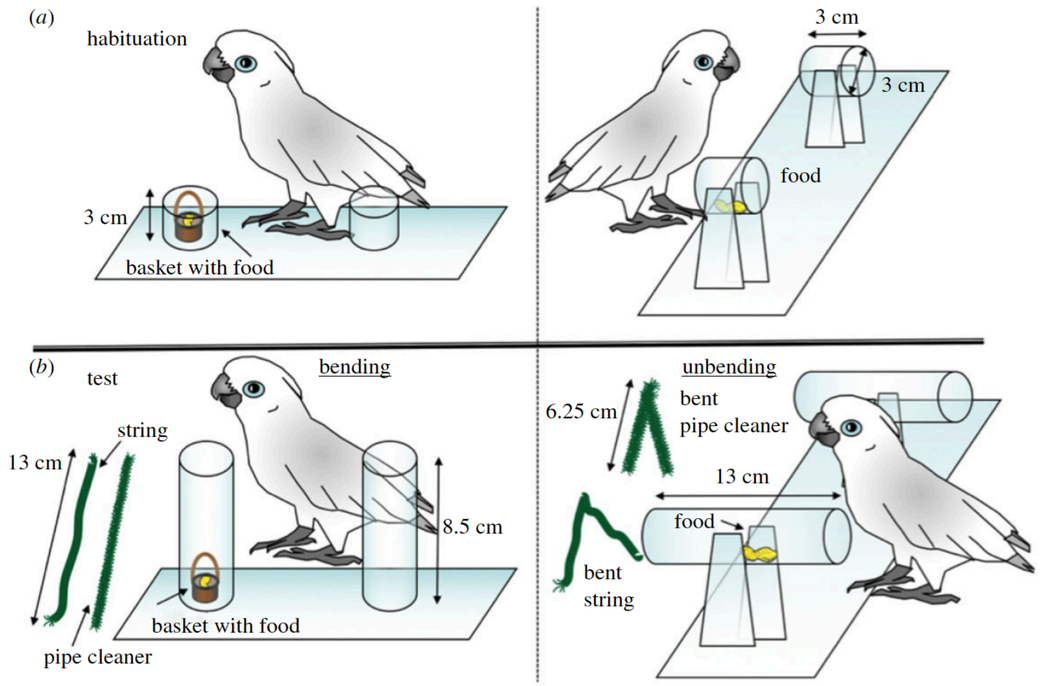


MY VIEW OF ANIMAL SELF-CONSCIOUSNESS

- A kind of pre-reflective, nonconceptual self-consciousness: part of the content of a conscious state that represents the body *as locus of perception and action* as *the same* across two 'temporal modes of presentation'. Model for this: an egocentric cognitive map, centred on 'here'.
- Its **function** is to allow the animal to keep track of the identity of its body across conscious representational states (thoughts) about present states of affairs, and states about possible states of affairs.
- This is likely necessary **when, and only when**, an individual subject must direct its own novel train of thought, as in individual innovation (a kind of instrumental practical reasoning).
- One situation in nature when an individual subject must individually innovate is when it solves practical problems for which the species has not already developed solutions (e.g., fairly rigid behavioural routines, perhaps 'coded in DNA' (c.f. Millikan (2004)).
- **Possible natural origin**: if a species faces such problems in its evolutionary history, then self-consciousness arises to facilitate its capacity for novel thought towards solutions.

EMPIRICAL CASE STUDY OF TOOL USE: LAUMER ET AL. 2017

- Laumer et al. (2017) designed experiments to test whether "innovative tool bending or unbending" occurs in a bird "that apparently lacks any ecological pre-dispositions for bending a material during tool-related foraging and nest building" (p. 2).



rspb.royalsocietypublishing.org Proc. R. Soc. B 284: 20171026

Figure 1. (*a*) Apparatuses for the habituation phase; (*b*) apparatuses for the testing phase. Left: vertical tube apparatus; right: horizontal tube apparatus. Dimension are given in centimetres.

EMPIRICAL CASE STUDY OF TOOL USE: LAUMER ET AL. 2017

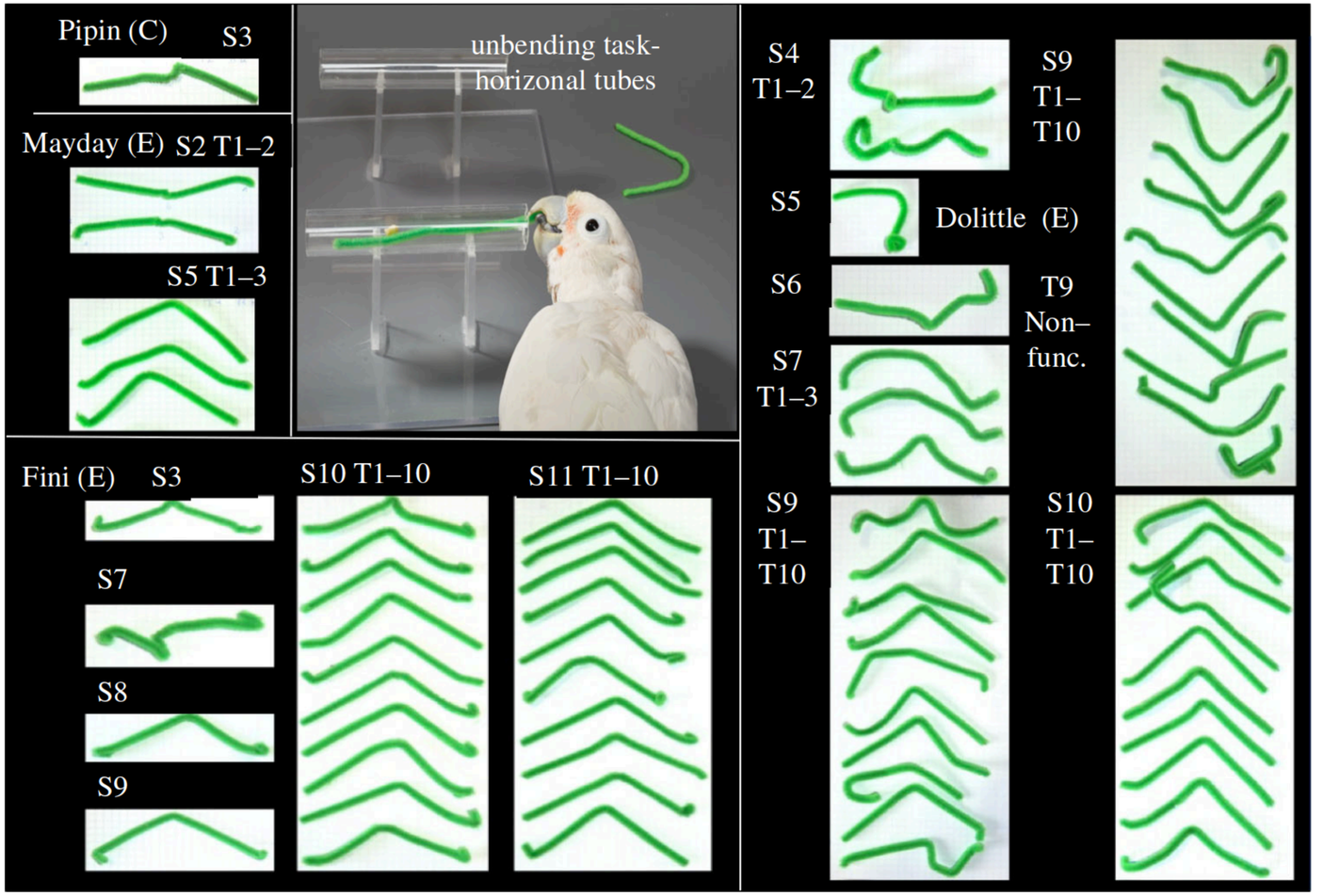
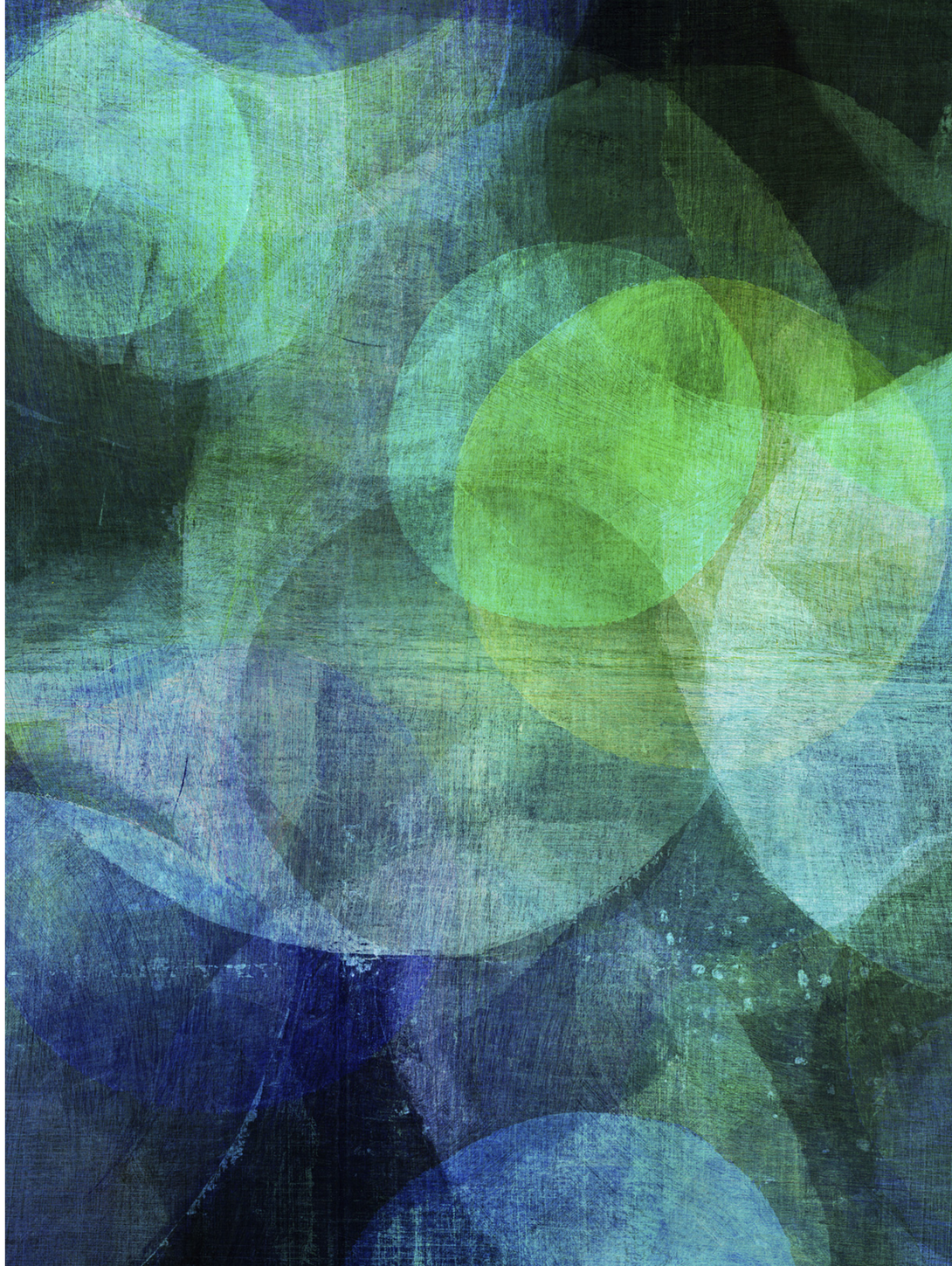


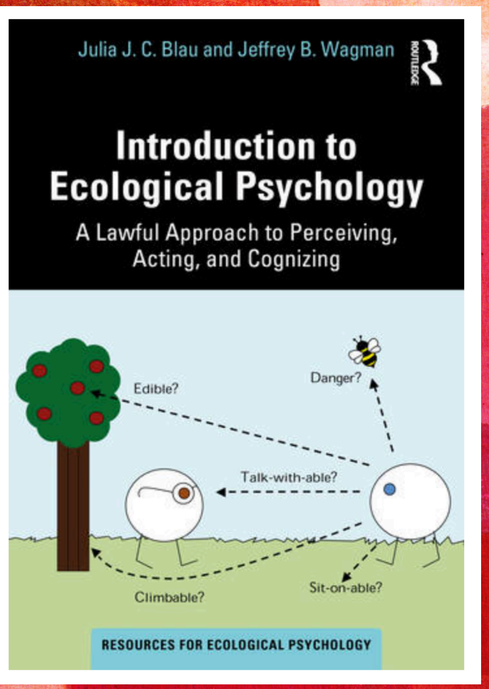
Figure 4. Successful stick-tools build by subjects Pipin, Mayday, Fini and Dolittle in the horizontal tube condition. S, session number; T, tool number; (C), control group/no experience; (E), experience group; Photograph by Bene Croy.

rspb.royalsocietypublishing.org *Proc. R. Soc. B* **284**: 20171026

SUMMARY OF HEGARTY (2025 & 2026)

- How do we explain the experimental observations?
- I spend considerable time ruling out simpler alternative explanations: including brute-force trial-and-error, ethological 'fixed action patterns', and biological functions selected for in evolution (teleofunctionalism).
- Key point: if the observed tool use is not species-typical behaviour (an empirically defeasible claim), then we need an explanation at the level of individual cognition.
- Proposal: likely the cockatoo thinks something to the effect of '**bending that thing this way makes it possible to get the food**'.
- So, they must be thinking of the pipe cleaner *as bent* (non-actual, prospective) and combining this representation with *current* goals and perceptions.





Egocentric cognitive map (no self-representation)



Content: ⟨to-be-avoided/climbed etc., __ in front of/behind/etc. [here]⟩

REPRESENTATIONAL COMBINATION

- Cognitively modest resources I draw on: (from ecological psychology/phenomenology/teleofunctionalism): the birds perceive 'affordances', or opportunities for action.
- Egocentric cognitive maps represent the world as affordances in the current environment from an egocentric perspective.
- Hegarty (2025) argues that non-actual states of affairs cannot appear on egocentric cognitive maps of the kind discussed by Schellenberg (2016). (c.f. Evans (1982); Alsmith (2000))

the creature's goal state. So, the explanation of the creature's behavior will attribute to it the following representational states:

(1) Goal state: desire for the food in cylinder.
(2) Perceptual states: sees the food in the cylinder, sees the pipe cleaner.
(3) $\mathfrak{R}^{\sharp}$: representation of the non-obtaining state of affairs *the pipe cleaner's being bent thus-and-so*.

Hegarty (2025), p.333



Here-and-now representation



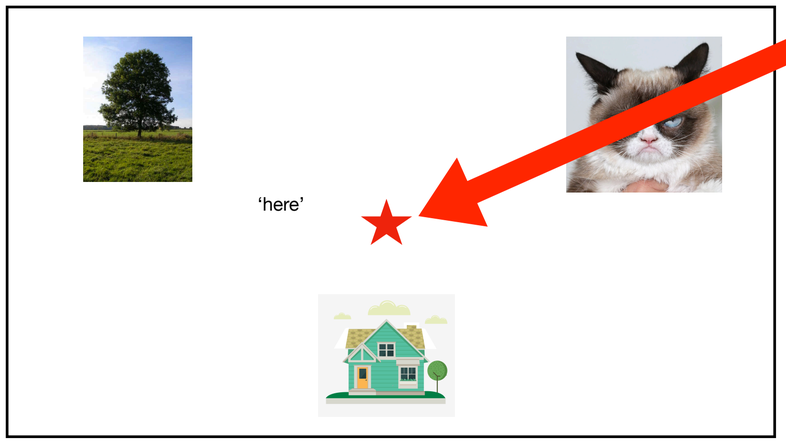
Prospective representation

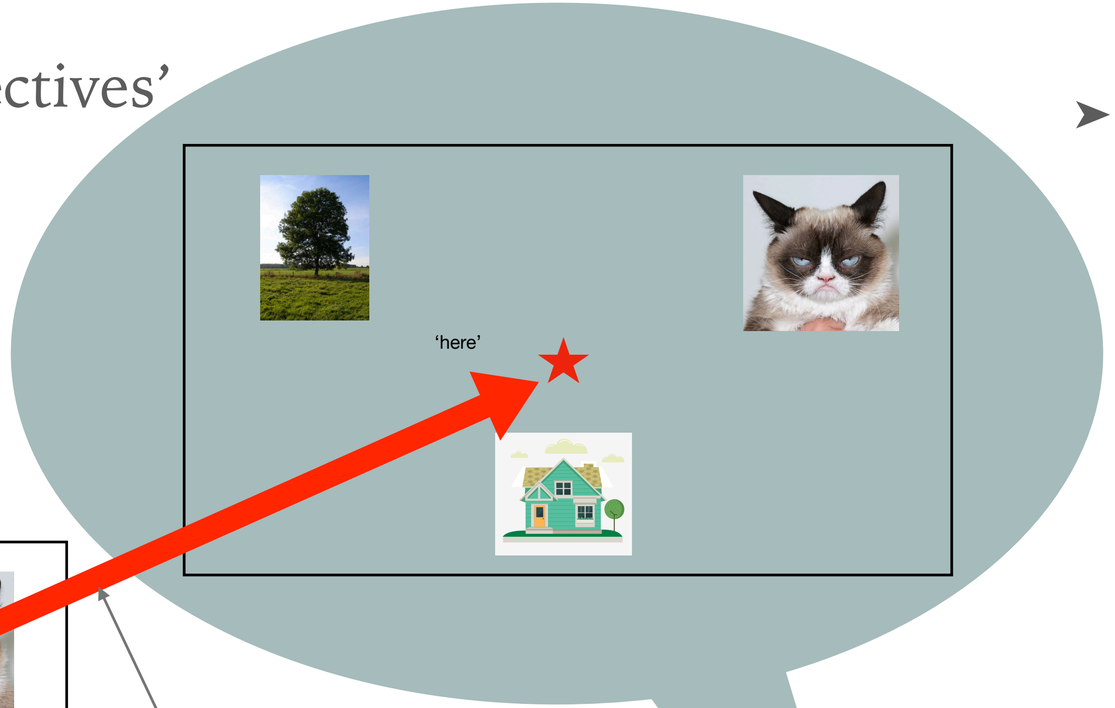
REPRESENTATIONAL COMBINATION

- Here-and-now: They see the affordance ‘to-be-bent’, and they can do so because their embodied capacities for interacting with the world structure what they see.
- Prospective: They run through this affordance ‘in their heads’ through ‘embodied imagination’, recognising that the outcome is a possible affordance that can extract the food.
- Precedent for embodied imagination of this sort, e.g., (Rucińska & Gallagher 2021). (In effect, the cockatoo is imagining *its own body* bending the pipe cleaner.)
- The cockatoo has to combine both representations; the same subject must be able to think them together (c.f. ‘access unity’ (Chalmers & Bayne (2010)). I suggest a common self-representational content makes this possible. (Why, more precisely? Coming up…)
- Egocentric cognitive maps represent opportunities for action in the current environment from an egocentric perspective. Prospective thought, even if structured as a map, could not be centred on the current environment.
- Hence, the ‘here’ is not the same. The content of self-representation identifies those ‘here’s as the **same** locus of perception and action. This, I claim, *is (deserves to be called) a kind of self-representation*.

Different ‘temporal perspectives’

Thought simultaneously





Self-representation ties together ‘body as locus of perception and action’ representations, i.e., identity of body, across temporal perspectives

Here-and-now

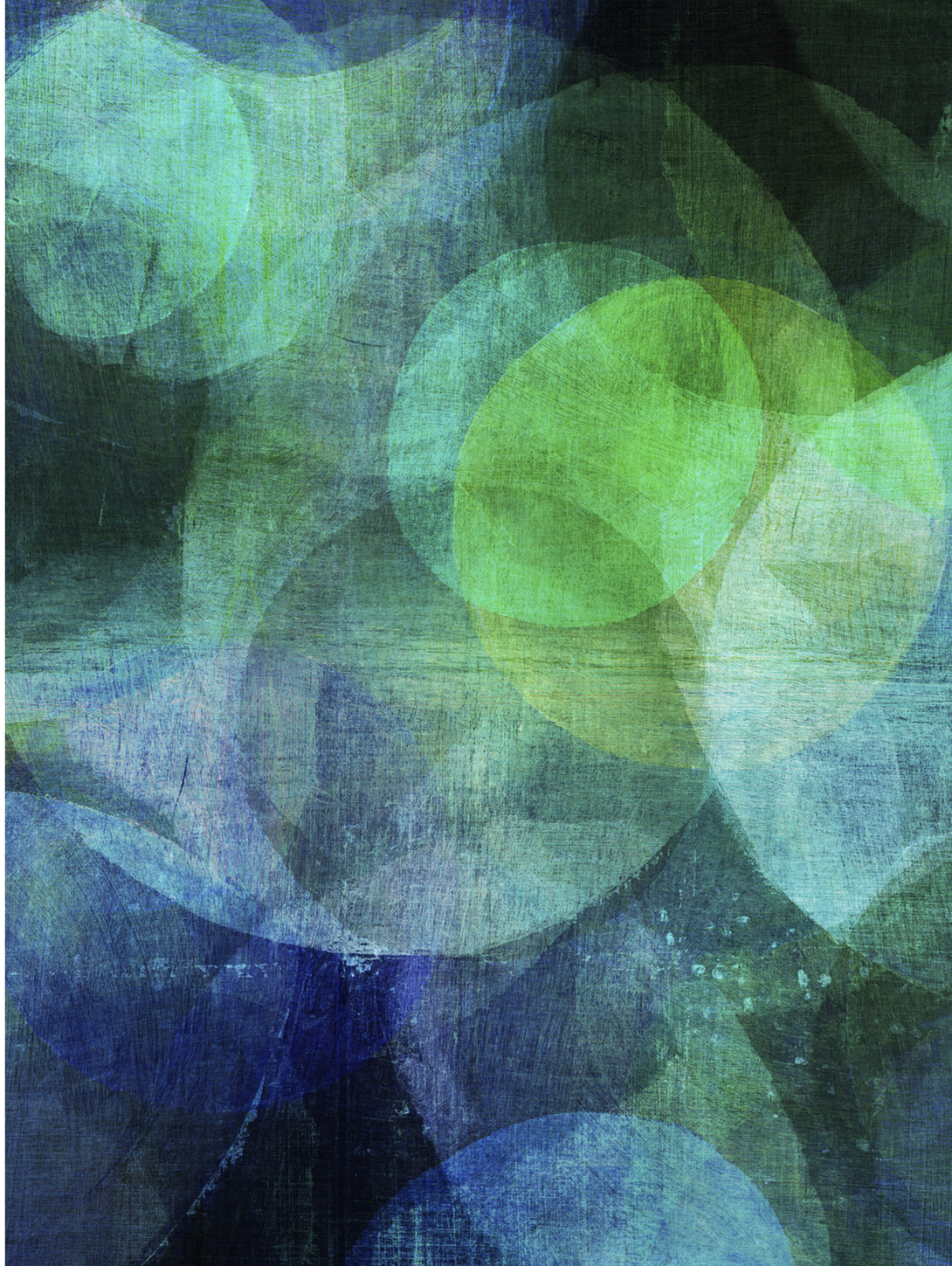
Prospective



- Given the embodied nature of the perceptual and imaginative capacities deployed, and the conscious and individual-directed nature of the cognition, the animal must recognise ‘current body’ and ‘imagined body’ as identical.
- Self-representation *just is* this representation of identity across temporal perspectives, I suggest.

CONCLUSIONS AND IMPLICATIONS

- I have used empirical and philosophical argument to describe, and motivate the need for, a practical representational capacity that is a kind of conscious self-representation.
- I have described a function for this capacity to play (unification of representations from different temporal perspectives). And this suggests a natural origin for self-consciousness in practical rationality of a certain sort (instrumental rationality involving thinker-directed ('active') cognition).
- This suggests that self-consciousness comes in varieties, or degrees.
- The full papers suggests that some animals have a degree of rationality, and a degree of stimulus-independence in thought.
- This makes progress on understanding **cognitive intermediates** between the simplest animals and fully rational human adults:
    1. cognition with partial conceptuality (and, hence, partial generality and recombinability),
    2. partial stimulus-independence,
    3. partial (domain-restricted) rationality, and
    4. partial (or some degree of) self-consciousness.



Thank you!

Email: michael.hegarty@oxon.org

Website: mjhegarty-philosophy.owlstown.net