Mistakes in action: on clarifying the phenomenon of goal-directedness

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Hill, J., Oderberg, D. S. ORCID: https://orcid.org/0000-0001-9585-0515, Austin, C., Cinotti, F. ORCID: https://orcid.org/0000-0003-2921-0901, Bojak, I. ORCID: https://orcid.org/0000-0003-1765-3502 and Gibbins, J. M. ORCID: https://orcid.org/0000-0002-0372-5352 (2025) Mistakes in action: on clarifying the phenomenon of goal-directedness. Biological Theory. ISSN 1555-5550 doi: 10.1007/s13752-025-00496-6 Available at https://centaur.reading.ac.uk/120955/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1007/s13752-025-00496-6

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Biological Theory
https://doi.org/10.1007/s13752-025-00496-6

ORIGINAL ARTICLE

Mistakes in Action: On Clarifying the Phenomenon of Goal-Directedness

Jonathan Hill[1,2] · **David S. Oderberg**[1] · **Christopher Austin**[1] · **François Cinotti**[3] · **Ingo Bojak**[4] · **Jonathan M. Gibbins**[3]

Received: 30 May 2024 / Accepted: 25 March 2025
© The Author(s) 2025

Abstract

Common sense tells us that biological systems are goal-directed, and yet the concept remains philosophically problematic. We propose a novel characterization of goal-directed activities as a basis for hypothesizing about and investigating explanatory mechanisms. We focus on survival goals such as providing adequate nutrition to body tissues, highlighting two key features—normativity and action. These are closely linked inasmuch as goal-directed actions must meet normative requirements such as that they occur when required and not at other times. We illustrate how goal-directed actions are initiated and terminated not by environmental features and goals themselves, but by *markers* for them. For example, timely blood clotting is the essential response to injury, but platelet activation, required for clotting, is initiated not by the injury itself but by a short sequence of amino acids (GPO) that provides a reliable marker for it. We then make the case that the operation of markers is a prerequisite for common biological phenomena such as mistake-proneness and mimicry. We go on to identify properties of markers in general, including those that are genetically determined and those that can be acquired through associative learning. Both provide the basis for matching actions to changing environments and hence adaptive goal-directedness. We describe how goal-directed activities such as bird nest construction and birdsong learning, completed in anticipation of actions in the environment, have to be evaluated and practiced against a standard of correctness. This characterization of goal-directedness is sufficiently detailed to provide a basis for the scientific study of mechanisms.

Keywords Action · Goal-directedness · Markers · Mistakes · Normativity

Introduction

Although there has been considerable philosophical skepticism regarding the reality of goal-directedness in biological systems, there is little doubt that there is a phenomenon deserving of further scrutiny. Birds accumulate sticks and stop when the accumulation is fit to protect eggs. Beavers assemble sticks and logs and stop when the assembly causes enough water to accumulate behind them. Squirrels store sufficient food to maintain them through the winter. Nevertheless, we still lack an adequate account of how to characterize these processes, and how to explain them.

In this article we lay out a novel perspective on goal-directedness. At its heart is the idea that normativity is integral to goal-directedness. Goal-directed activities can proceed effectively, appropriately, successfully, correctly, or else ineffectively, inappropriately, unsuccessfully, or mistakenly. These are all essentially normative terms pointing to how *well* or *badly* a living thing gets on in its environment. We consider normativity in the most general of terms. We do not link it conceptually to intention, consciousness (let alone self-consciousness), language, free will, "value judgments," obligations, or any of the other features of normativity that give it shape in the human species. We do not

✉ Jonathan Hill
Jonathan.hill@mmu.ac.uk

[1] Department of Philosophy, University of Reading, Reading, UK

[2] School of Psychology, Manchester Metropolitan University, Manchester, UK

[3] School of Biological Sciences, University of Reading, Reading, UK

[4] School for Psychology and Clinical Language Sciences, University of Reading, Reading, UK

Published online: 18 April 2025

enter into technical debates about whether normativity is "naturalistic" or "non-naturalistic." All we claim is that normativity, in the context of a theory of biological mistakes, is naturalistically respectable in the sense of being observable, testable, and a fit subject for empirical research.

The normative demarcation between the correct and the incorrect is determined by the potential for effective action in the environment (Hill et al. 2022; Oderberg et al. 2023, 2024). Actions are behaviors directed towards objects or events in the environment of the system, many but not all of which are essential to survival and reproduction, such as potential mates, food, shelter, and predators. We argue that once we recognize the centrality of normativity and action, we are in a position to spell out the further details of goal-directed actions. Our method aims to build on the tradition exemplified by C.H. Waddington ([1957]2014) of describing the nature of goal-directed activities without preconceptions about the mechanisms and processes involved.

At the same time, we aim to operationalize our account of goal-directedness by showing its direct relevance to biological research. We do this first by examining hemostasis (blood clotting), showing what its characterization looks like both prior to and after making explicit the place of action and normativity. We then look in detail at the way in which the activation of clotting is linked to blood-vessel injury, highlighting that it is not the injury itself but a *marker* for it which is the causal agent. A marker, as we describe it, is a physical entity that provides an indication of the presence of environmental objects or events requiring, or at least potentially requiring, goal-directed action. Markers can also be identified that are indicators of whether the conditions for the termination of action have been reached. Markers are subject to the normative tests of whether they provide the basis for reliable, accurate, timely, and effective action.

We go on to illustrate how our characterization of goal-directedness is operationalized in two further examples: satiety and the termination of feeding; and identification and location of sources of food. We then move from particular analyses to a broader justification of the role of markers by pointing out how phenomena that are widespread throughout biological systems, mistakes and mimicry, are only possible because of resemblances between markers for two or more different environmental events or objects. We also bring out a remarkable property of markers—that their physical structures are arbitrary inasmuch as they can vary substantially provided they meet the relevant normative requirements. This property is also at the heart of associative learning whereby new markers are established as the result of experience, starting with neutral objects. This provides the basis for flexible, adaptive, goal-directed actions in variable or unpredictable environments. Finally, we analyze goal-directedness where action directed towards an intermediate goal, such as nest construction, must be completed *prior* to the realization of the goal, in this case, successful fledging of chicks. This highlights that the organism has to work towards an internal standard for the goal and that mistakes are now variations that fall short of the standard.

Background

The history and scope of theories of goal-directedness have been the focus of several recent reviews and critiques (Okasha 2022; Dresow and Love 2023). These important surveys highlight how diverse the theories are and show that, in many respects, they can be somewhat vague. According to Dresow and Love (2023, p. 109), "Currently there is no consensus on what 'goal-directedness' means in biology." Okasha has posed the questions of whether the concept is anything more than a superficial characterization of resemblances across diverse biological processes, or perhaps merely a construct imposed by our human inclination to see purpose in the natural world.

A key question, and one of central importance to the main proposal of this article, is whether goal-directedness should be characterized by what it is or by what explains it. Dresow and Love draw attention to this distinction in their discussion of teleology, contrasting the "phenomenon to be explained" and the "form of explanation" (2023, pp. 108–109). The introduction of the term "teleonomy" to denote the phenomenon shorn of what some have seen as its metaphysical baggage may appear scientifically more acceptable, but its contribution remains in doubt.

Many accounts of what the phenomenon is converge on the idea that there are activities in biology that show "end-directedness" (Pittendrigh 1958, p. 394). Waddington ([1957]2014, p. 2) put this succinctly: "most of the activities of a living organism are of such a kind that they tend to produce a certain characteristic end result." This characterization invokes two distinct properties—directedness and termination. By contrast, behaviorist accounts of what goal-directedness is have sought to characterize it in terms of persistence and plasticity (Rosenblueth et al. 1943; Braithwaite 1953; Lee and McShea 2020).

Explanatory accounts of goal-directedness commonly refer to processes akin to intentions, beliefs, or desires in the way we understand them in humans (Nissen 1997), or to a genetic program (Mayr 1988). Despite the distinction between the phenomenon itself and its possible explanation, discussions of goals still commonly conflate the two, for example by using the terms "goal-directed" and "having a goal" interchangeably. Setting aside the question of how to operationalize "having a goal," this conflation implies that in order to be goal-directed the system or organism must *have* a goal. The aim of this article is to characterize the phenomenon of goal-directedness independently of whether this necessitates having a goal, or an intention, indeed independently of any other explanatory mechanism.

Defining goal-directedness in terms of putative explanations is, we suggest, problematic in two ways. First, without establishing what requires explanation it is difficult, perhaps impossible, to evaluate the status of candidate explanations. We will consider examples of various kinds of behavior, including sub-organismic systemic behavior, in order to evaluate whether the same analysis of goal-directedness applies in the same way across all of them. We include the sub-organismic system of blood clotting, a system that has not appeared in previous accounts of goal-directedness and might on some perspectives be excluded. We are not alone in encompassing a variety of biological phenomena. For example, Lee and McShea (2020) refer to bacterial and human behaviors in their analysis of the phenomenon. This suggests that an adequate characterization of goal-directedness as such is required in order to clarify what should be included and what excluded, *prior* to an exploration of possible mechanisms.

Secondly, conflating the phenomenon with its possible explanations runs the risk of not separating the philosophical from the empirical questions. We illustrate this broad concern by an examination of bird nest construction and song learning. Bird nests are complex structures; however, studies have considered the possibility that this complexity may only require simple rules, either for a predetermined stereotypical series of behaviors, or based on feedback from local nest cues (Walsh et al. 2013). Alternatively, more complex cognitive processes involving memory and learning may be required (Lehtonen et al. 2023). Any philosophical argument requiring one or the other kind of mechanism for behavior to qualify as goal-directed, risks preempting the scientific enquiry. Similarly, in contrast to Mayr's proposal that goal-directedness entails a "genetic programme" (Mayr 1988), empirical investigations of bird song find that in some species song performances are directed by a genetic program and in others they are learnt. If Mayr's requirement based on philosophical grounds were to be accepted, this would seem to rule out alternatives that should remain open to scientific investigation.

The problem of conflating the phenomenon with its potential explanation arises in relation to some formulations of the role of agency in goal-directedness. Put simply, agency refers to the capacity for action, and actions are commonly taken to be the result of a certain "psychological architecture" (Okasha 2018, p. 13) in which there are belief-like or desire-like states (Schlosser 2015). This clearly belongs with explanatory accounts. Suppose then that goal-directed phenomena of the kind characterized by Waddington are not explained by such belief- or desire-like states: would they no longer be considered goal-directed? This would seem to be problematic. Similarly, if we assume that these states can be operationalized for scientific investigation, there would be a strong case for studies that are open to the possibility of belief-like or desire-like states among a number of possible mechanisms, just as in the scientific investigations of bird nest construction.

There is a common presumption that kinds of behavior formulated in terms of belief-like or desire-like states are the paradigm cases of biological agency. The alternative is to see them as instances of something more generic, such as a Minimal Biological Agent (Oderberg et al. 2023). Looked at in this way, the paradigm cases might be found in biological systems that display Waddington-type behaviors but indisputably do not have belief-like or desire-like states. This observation underpins our choice of blood clotting in response to tissue injury, and satiety behaviors in the termination of eating, as exemplars of the phenomena we aim to elucidate. Agency underpinned by beliefs or desires, or states resembling them, would then comprise particular, complex, or elaborated examples of a much broader set of phenomena in biology. In that case goal-directedness would have a "behavioral architecture" with a connection—seamless in some ways, substantially different in others—to human "psychological architecture." The characterization of goal-directedness that we seek to identify in this article is rooted in an analysis of biological systems in general, without reference to concepts borrowed or extended from human mental activities.

Goal-Directedness: Normativity and Action

Nagel (1979) similarly argues against the extension of concepts drawn from human mental activities, and in favor of seeking a formulation of goal-directedness that would not look out of place in general biology. At several points he comes close to identifying the role of normativity, referring for example to the possibility of goal-directed behaviors being mistaken and in need of correction, without however providing an analysis of what "mistaken" and "corrected" mean here, nor placing them as we do at the heart of the analysis. Nevertheless, and crucially for our purposes, he concludes a process of argument and counterargument with a precisely operationalized formulation (Nagel 1979, pp. 311–312):

> **G_0**: "a function of item *i* in system *S* and environment *E* is *F*, presupposes…that *S* is goal-directed to *some*

> goal G, to the realization or maintenance of which F contributes."

We will use this formulation as a fruitful starting point on which to scaffold our normative analysis of biological goal-directedness.

Before taking the analysis forward using Nagel's characterization, note the way he links the concept of *function* to the idea of goal-directedness by presupposing that the function ascribed to an item *contributes* to goal-directedness. In this formulation "function" and "goal-directedness" are tightly connected. This use of the concept of function as biological role is consistent with the way it is handled in biological mistake theory, where it broadly means "what the item does in respect of" or "how the item contributes to" some behavior of the system (see further, Bolton and Hill 2004; Hill et al. 2022; Oderberg et al. 2024). This concept of function coincides with the way it is applied in biomedical research, meaning that the test of whether item i has function F is whether the outcome of a biological process is impacted by alterations in, or by the absence of, i. By contrast it does not coincide with the concept as it occurs in questions such as "what is the function of i?" where that is generally taken to be a question about why the item is there in the first place. The answer to this question is generally assumed to invoke its evolutionary origins, as per the popular "selected effects" theory of function whereby a function is identified by its evolutionary history (e.g., Neander 1991; Garson 2019).

Biological systems, organisms, or collectives must act effectively in their environments: to move away from danger, to respond to threats to the integrity of the body, to move towards food and to feed, and to regulate key parameters such as body temperature and oxygen supply to tissues. These are all essential for survival. Note, however, that not every goal is a survival goal in the specific sense of *contributing directly to life extension as an overarching final goal*—self-sacrificial behavior being a case in point (though here one could advert to survival goals for the colony, etc.). For our case, however, there only needs to be agreement that there are survival goals in the specific sense of *keeping the organism alive*, such as obtaining adequate nutrition, avoiding predators, maintaining body temperature, fighting infections, and reproducing. We are not aware of an argument against this claim, which provides a starting point for our analysis of goal-directedness.

Accounts of goal-directedness, however, rarely draw out the specific parameters of action that are essential to these outcomes. These parameters include not only what the action consists in, but also the conditions under which it is initiated and terminated, and to what features of the system's or organism's environment the action is oriented. Based on these considerations we propose that actions, conceived of in this way, are integral to many instances of goal-directedness. This is implicit in Nagel's formulation and is readily made explicit by the additions in bold:

> $\mathbf{G_1}$:a function of item i in system S and environment E is F, presupposes that **action generated by** S is directed to some goal G to the realization or maintenance of which F contributes.

The importance of such parameters becomes clear when we consider normativity together with action. Biological actions are subject to normative criteria.[1] Systems and organisms have to act reliably and accurately and must be timely and effective in their environments—blood clots to stem blood loss, immune systems to combat infections, bees to locate nectar, beavers to build dams, spiders to spin webs, and birds to build nests. The reliability criterion refers to acting in this way repeatedly over time; accuracy refers to acting in the right way under the right circumstances; timeliness refers to actions being neither premature nor too late; and effectiveness refers to whether the action is successful in relation to environmental opportunity such as food, or risk such as predator. The extent to which actions are mistake prone highlights their normativity. For example, blood clots—which we will discuss in more detail later—may occur when they should not, blocking arteries and causing cardiovascular disease. They can also fail to occur when they should, leading to life-threatening bleeding. Many other parameters of clotting are open to the possibility of mistakes: initiating too late, terminating too late or too early, and so on. Turning this around and generalizing, actions have to occur when required, to not occur when not required, be timely, initiate and terminate at the right time, be large or strong enough, but not too large or strong, and must be directed towards relevant environmental objects or events.

We now introduce the normativity of goal-directedness into Nagel's characterization shown in bold italics:

> $\mathbf{G_2}$: a function of item i in system S and environment E is ***appropriate-and-timely-F***, presupposes that

---

[1] The case for the centrality of normativity is made in the context of a longstanding debate about whether the normative is "nonnatural" (e.g., Davies 2003; Wedgwood 2007; Thomson 2008). Our arguments for the universality of normativity in biological systems, focusing on the contrast between correctness and mistakes, can be found in Hill et al. (2022) and Oderberg et al. (2023, 2024). The mistake theorist sees normativity as a real and ubiquitous phenomenon across biological systems and amenable to scientific investigation, hence philosophically and scientifically naturalistic. That said, in the context of the present analysis the debate about whether normativity is in some sense also "nonnatural" can be placed to one side.

> ***accurate and effective*** **action** by that *S* is directed ***accurately*** to ***relevant*** goal *G* to the realization or maintenance of which ***appropriate-and-timely-F*** contributes.

Normativity, Correctness/Mistakes and Goal-Directedness: Three Case Studies

Blood Clotting

Blood clots form in arteries in response to arterial wall injury. The goal of clotting is to prevent life-threatening blood loss. Activation of platelets is a key step in the initiation of clotting, and receptors on platelets, including the glycoprotein GPVI, are necessary structures in this sequence (Gibbins 2004). We can apply this example to Nagel's characterization, with references to clotting mechanisms in italics:

> $\mathbf{G_3}$: a function of item *i* (*the glycoprotein (GPVI) receptor on platelets)* in system *S* (*hemostatic*) and environment *E (sharp objects*, *injury)* is *F* (*initiation of clotting*), presupposes that *S (hemostatic*) is directed to goal *G* (*preventing life-threatening blood loss*) to the realization or maintenance of which *F* (*initiation of clotting*) contributes.

Then making explicit the place of action:

> $\mathbf{G_4}$: a function of item *i* (*the glycoprotein (GPVI) receptor on platelets) in system S (hemostatic*) and environment *E* (*sharp objects*, *injury*) is *F* (*initiation of clotting*), presupposes that **the action** *(clot formation in the blood vessel)* **generated** by *S* (*hemostatic*) is directed to goal *G* (*preventing life-threatening blood loss*) to the realization or maintenance of which *F* (*initiation of clotting*) contributes.

And then making the normativity explicit in bold italics gives us:

> $\mathbf{G_5}$: a function of item *i* (*the glycoprotein (GPVI) receptor on platelets*) in system *S* (*hemostatic*) and environment *E* (*sharp objects*, *injury*) is *F* (***appropriate and timely*** initiation of clotting), presupposes that ***accurate and effective*** **action** ***(well placed***, ***neither too small nor too large*** *clot in the blood vessel)* **generated** by *S* (*hemostatic*) is directed to goal *G* (*preventing life-threatening blood loss*) to the realization or maintenance of which *F* (***appropriate and timely*** *initiation of clotting*) contributes.

Goal-directed effective action commonly requires accurate and timely responses to objects or events in the environment of the system or organism. How is this achieved? The activation of platelets and hence the initiation of clotting occur in response to blood-vessel injury, but platelets are not activated by injury and there are no "injury receptors." Rather, platelets are activated by a large protein, collagen, to which they are reliably exposed when there is an injury; not however by the entire collagen molecule but by a small subset of amino acids on it, several repeats of glycine-proline-hydroxyproline (GPO) triplets. So our formulation now reads (with the normative qualifiers omitted for clarity and the marker in bold):

> $\mathbf{G_6}$: a function of *i* (*the glycoprotein (GPVI) receptor on platelets*) in system *S* (*hemostatic*) and environment *E* (*sharp objects*, *injury associated with* ***GPO triplet exposure)*** is *F* (*initiation of clotting*), presupposes that the action *(clot in the blood vessel)* generated by *S (hemostatic*) is directed to goal *G* (*preventing life-threatening blood loss*) to the realization or maintenance of which *F* (*initiation of clotting*) contributes.

Given the priority for rapid action following injury there seem to be several advantages of a sequence of this kind—injury followed by collagen exposure and then initiation by GPO exposure on the collagen molecules—compared to initiation of clotting by the injury per se. Injuries can differ in many ways, which would necessitate the identification and appraisal of many complex and varying environmental objects by the clotting system before effective action was possible. Collagen consists of large molecules comprising several thousand amino acids, of which variable proportions would be likely exposed following injury, again offering a vast range of objects for appraisal prior to initiation of clotting. A small subset of amino acids, by contrast, presents a relatively simple and constant object for appraisal, one which is readily detected at a receptor site.

This brings us to a key point. The injury, to which clotting is a response, does not initiate the clotting. That requires exposure to GPO triplets. Put more generally, the environmental event or object to which the goal-directed behavior is a response does not initiate the behavior. That requires a *marker* for it. By "marker" we mean a physical entity that provides an indication of the presence of environmental objects or events requiring, or potentially requiring, action.[2] The necessary feature of a marker in blood clotting is that it has a systematic association with the key event or object

---

[2] The term "marker" in this context is used in the biomedical literature, albeit without the kind of technical account we are offering. See, e.g., https://my.clevelandclinic.org/health/diseases/24067-antigen (last accessed 15 March 2025).

and that this ensures reliable, accurate, and timely clot initiation. It must meet a standard of correctness since it has to ensure reliable, accurate, and timely clot initiation. The normative requirements which we described for clotting apply equally to the GPO amino acid triplets themselves. It is essential that the GPO amino acid triplet marker does not commonly trigger platelet activation in the absence of injury, and conversely that it reliably triggers a timely sequence where there is an injury. The role of markers for environmental objects and events is central to our account of actions directed towards survival goals, as will be seen.

Termination of Feeding

Our next step towards justifying the view that these properties of goal-directedness are common across biology is to consider a different normative parameter of action in a wholly different context: termination of feeding. Termination of feeding is coordinated by brain activity across many species. We therefore expect brain-directed termination of feeding to be linked systematically to its goal—the provision of appropriate levels of nutrition to tissues. This is achieved via feedback to the brain regarding the progression of the behaviors towards the goal. Using Nagel's formulation again, this can be summarized as follows with the normativity in bold italics:

> **G$_7$**: a function of item *i* (*electrical impulses in afferent nerves*) in system *S* (*feeding and satiety)* and environment *E* (*presence of food*) is *F* (***timely termination*** *of feeding*), presupposes that termination of the action *(feeding*) generated by *S* (*feeding and satiety*) is directed to goal *G* (***maintenance of the right levels of nutrition to vital organs***) to the realization of which *F* (***timely termination of feeding***) contributes.

The challenge of timeliness is again paramount. The goal is broad and complex, involving a number of nutrients such as glucose, lipids, and amino acids supplied to many organs and parts. Could the action—feeding—be terminated by achievement of the goal of delivery of adequate nutrition to tissues? Crucially for timeliness, food must be digested, cross the gut wall, and be delivered via the bloodstream to tissues in order for the goal to be achieved. This process varies by nutrient but commonly takes several hours. If feeding terminated only upon attainment of the goal of providing adequate tissue nutrients after several hours, this would result in gross overfeeding and ultimately the overwhelming of tissue metabolism (Watts et al. 2022).

This brings us to the same key point already made in respect of clotting. The internal state that leads to termination of goal-directed behavior—in this case feeding—is not the achievement of the goal—adequate nutrition to the tissues—but a *marker* or in this case *multiple markers* for its provision. The markers, individually or acting in concert, must meet a standard of correctness based on whether they ensure reliable, accurate, timely, and effective termination of feeding. In mammals we find that markers contributing to termination of feeding include several hormones as well as distension of the stomach and small intestine (Owyang and Heldsinger 2011; Watts et al. 2022). For this analysis we focus on one marker, gastric distension.

Hence we can reformulate the goal-directed sequence highlighting the role of a termination marker in bold, and omitting the normative qualifiers, along Nagelian lines as follows:

> **G$_8$**: a function of item *i* (*electrical impulses in afferent nerves*) in system *S* (*feeding and satiety)* and environment *E* (presence of food associated with ***gastric distension***) is *F* (*termination of feeding*), presupposes that termination of the action *(feeding*) generated by *S* (*feeding and satiety*) is directed to goal *G* (*maintenance of nutrition to vital organs*) to the realization of which *F* (*termination of feeding*) contributes.

In many respects this could be seen as a paradigmatic example of goal-directedness, in which the behavior is regulated according to cybernetic principles by feedback to the organism from the effect of the action in relation to its goal (Rosenblueth et al. 1943). We make two points in relation to this example. First, feedback via neural mechanisms provides a rapid and efficient way of tracking the progress of an action, but our analysis does not assume that feedback mechanisms require neuronal mediation. Secondly, neither do we assume that mechanisms for termination of action require feedback on their effectiveness with respect to the goal. For example, while this has not been ruled out for termination of clotting, it seems unlikely that this is the kind of mechanism involved.

Anticipation of Accurate and Timely Action in an Environment of Varying Food Provision

Animals in their natural habitats typically have to identify which objects are nutritious, contrasted with those that are not (whether neutral or harmful), and to locate and approach them. This introduces additional requirements for the normativity of action. Given that feedback on the food status of the environmental object would be hazardous once feeding has started, anticipatory mechanisms are required. Furthermore, in a complex and changing environment, for example where locations of food change and sources of threat are mobile, an organism must be able to predict future states and events such as the presence of food and of danger. Predictions create time in which actions can be matched to environments, thus avoiding the mistakes of losing resources or of falling victim to danger. (Schultz et al. 1997; Chen et al. 2015; Morand-Ferron 2017).

Action is based on responses to markers, such as color, shape, or odor, in anticipation of the likelihood that the object will turn out to be nutritious or dangerous. So we encounter again the phenomenon that we identified in relation to the initiation of clotting—that the goal-directed action must be based, not on the environmental object to which the behavior is directed, but on a reliable marker for it. Our Nagelian functional statement, highlighting the food odor marker and the normative qualifiers in bold, can be written as follows:

> **G$_9$**: a function of item *i* (*odor receptor/sensor activations*) in system *S* (*feeding*) and environment *E* (*presence of food associated with **odor***) is *F* (***accurate** food identification and location*), presupposes that ***accurate and effective*** action ***(timely**, **well-directed** approach)* generated by *S* (*feeding*) is directed ***effectively*** to goal *G* (*ensuring **adequate** nutrition*) to the realization or maintenance of which *F* (***accurate** food identification and location*) contributes.

How then are markers such as odor linked to likely nutritional value? Evaluation of the likely nutritional value of an object, of anticipatory behaviors such as food seeking, and of feeding behaviors, are regulated by reward mechanisms that are broadly characterized in terms of "liking" and "wanting" (Morales and Berridge 2020). Far from being anthropomorphic extensions of typically human motivational states, extensive evidence from across many species identifies in more general and wide-ranging terms these two components of reward mechanisms and their neuronal circuitry. More formally, "liking" refers to the hedonic value of a food or food cue, which in animals is assessed by characteristic orofacial expressions (Morales and Berridge 2020, p. 2). "Wanting" refers to giving priority to the pursuit of a food or cue. Although these are complex, diverse phenomena, a considerable body of research converges on the role of the neurotransmitter, dopamine (Schulz et al. 1997). In studies in which the activity of single neurons in the midbrain is recorded in alert monkeys, the actions of touching fruit or tasting fruit juice—involving two anticipatory markers for nutritional value—are associated with increased activity of dopaminergic neurons. We can now redraft the functional statement for dopaminergic activity as follows:

> **G$_{10}$**: a function of item *i* (*odor receptor/sensor activations and associated increased dopaminergic activity*) in system *S* (*feeding*) and environment *E* (*presence of food associated with **odor***) is *F* (***accurate** food identification and location*), presupposes that ***accurate and effective*** action ***(timely**, **well-directed** approach)* generated by *S* (*feeding*) is directed ***effectively*** to goal *G* (*ensuring **adequate** nutrition*) to the realization or maintenance of which *F* (***accurate** food identification and location*) contributes.

Markers: The General Case

So far we have presented an account of the way markers for environmental objects or events can be observed in biological systems. In this section we take four further steps. First, we provide a formal definition of a marker. Secondly, we show how markers are necessary for the normative behavior and properties of biological systems. Thirdly, we argue that the arbitrariness of markers is key both to correctness and mistakes in environment-dependent goal-directedness. Fourth, we argue that this arbitrariness is central also to the acquisition of new markers and hence to learning.

Markers: A Formal Definition

As we noted earlier, the environmental event that leads to initiation of goal-directed behavior in clotting is neither the injury nor blood loss, but a *marker* for these. By marker we mean a physical entity that provides an indication of the presence of environmental objects or events requiring, or at least potentially requiring, goal-directed action. The central requirement for a marker in blood clotting is that it has a systematic association with the key event or object and that this ensures reliable, accurate, and timely clot initiation. It must meet a standard of correctness according to whether it ensures reliable, accurate, and timely clot initiation and termination. The normative requirements which we described for clotting apply equally to the GPO amino acid triplets themselves. It is essential that the GPO amino acid triplet marker does not commonly trigger platelet activation in the absence of injury, and conversely that it reliably triggers a timely sequence where there is an injury.

More technically, we can define a marker as follows:

Def. Marker

x is a marker for y in respect of organism O and goal G in environment E=def:

1. y is an entity towards which O's behavior is G-directed;
2. $x \neq y$;
3. Either:

   3.1 x is a feature of y; or

3.2 x is wholly distinct from y and is systematically associated with the existence of y; and

4. x could be used by O to achieve G.

Let us unpack this definition. The four entities involved in the concept of a marker are: the marker itself; what it is a marker for (i.e., the environmental object or event—let us call it "object" for brevity); the organism for which the marker has usefulness; and the environment in which the other three things exist—within which the organism undertakes, or could undertake, effective action to achieve the goal.

For clarification "organism" is used as a shorthand to cover not just individual organisms—the paradigm entities that behave goal-directedly—but also collectives such as species, populations, colonies, families, and so on, as well as *parts* of organisms and *sub-systems*, such as the hemostatic system, belonging to organisms. It includes also entities within organisms and their subsystems, such as antibodies, individual platelets, hormones, and even symbiotic ones such as gut microbes.

The above definition states that the marker is *distinct* from environmental object or event that constitutes "the entity towards which O's behavior is G-directed." Call this the "object" for short. It will be distinct in one of two ways. Either it is a *feature* of that object, such as a part, a property, a combination of the above, or even what we might call an *aspect* of the object. By aspect we mean the way the object *presents* itself to the organism, whether visually, auditorily, or via some other sense modality or combination of modalities. The other way of being distinct is being *wholly* distinct, by which we mean that the marker is not a feature of the object but some entity that is spatiotemporally (or spatially alone, or temporally alone) distinct from the object.

Now, a feature or part of an object is of course systematically linked to it. Where you find hearts you find valves; where you find rabbits you find rabbit parts (to borrow W. V. Quine's famous example (1960)). But a wholly distinct marker also needs to be *systematically associated* with the object for the sake of the organism's goal-directed behavior. Systematic association is a necessary condition of the organism's being able to *use* the marker to achieve the goal (as per clause 4). In other words, the marker needs to be *useful* to the organism, in a given environment, for generating effective goal-directed behavior. We discuss this further in the next subsection in relation to Millikan's consideration of representational content (Millikan 2024).

Markers: Foundational for the Normativity of Biological Systems

We focus here on the related phenomena of mistakes and mimicry. Examples of mistakes include: clotting in response to GPO triplets on an atheromatous plaque in the absence of an injury, resulting in arterial blockage; immune activation directed towards self rather than foreign antigens, resulting in autoimmune disease; infectious agents evading immune responses by resembling host cells or molecules, hence increasing their infectivity. We refer to these as mistakes of *misidentification* (Hill et al. 2022).

Examples of mimicry include: parasitic bird young singing songs that resemble those of the biological offspring of the parent birds, hence being fed by them; "bleeding flowers" resembling wounded insects, hence attracting fly pollinators; the Venus fly trap emitting an odor that resembles food odors, attracting the fly prey to the trap.

All these phenomena are possible because the system, such as clotting or immune, or the organisms, such as bacteria, birds, or flies, respond to *markers* for the relevant objects rather than the objects themselves. Platelets are mistakenly activated by repeat GPO triplets that are markers for injury even if the triplets are not generated by tissue injury. Immune systems are mistakenly activated by molecules that are markers for infection even when they are presented by body cells and not by an infectious agent. They also mistakenly fail to be activated when infectious agents present molecules that look like those of body cells. Adult birds respond to the mimicry of the song markers for young of their own species and feed them. And flies respond to the mimicry of visual markers for injured organisms and odors associated with food even in the absence of such objects.

Markers: Correctness and Mistakes

A remarkable and important property of markers is that although the physical structure and causal role of each marker can be described, a different structure may be an equally effective marker, provided it ensures the same reliable, accurate, and timely goal-directed action in the environment. Returning to platelet activation and clotting, if we label the repeat GPO triplets Marker 1 (M1), then the response to some variation on this, M2, will be equally correct if it ensures reliable, accurate, timely, and effective initiation of goal-directed behavior, while a response to a different variation, M3, will be a mistake if it *impairs* reliable, accurate, and timely initiation. We can also envisage a further variation, M4, which has an advantage over M1: suppose it leads to faster platelet activation and hence more effective protection from blood loss. Note that we cannot get the M2 – M3 – M4 contrast from a purely *physical* description of their differences from M1. The physical difference between M1 and M2 may be substantial without significantly affecting action—the action will still be effective—while the physical difference between M1 and M3 may be minor, yet the implications for action significantly different: the response to M1 may be lifesaving but the response to M3 fatal. Similarly, the physical difference between M1 and M4 may be minor while M4 nevertheless confers a substantial advantage in terms of life or health.

Note further that the status of responses to each of M1, M2, M3, and M4 as correct or mistaken may change under different environmental conditions. For example, if conditions were to arise that increased the risk of clotting, threatening to block arteries, then responses to M1, M2, or M4 could cease to be correct, while a response to M3, by reducing the likelihood of clotting, would become advantageous. In the case where M3 is the result of a heritable mutation it may confer selective advantage under these altered conditions. We can apply the same analysis to our account of markers for termination of feeding and the presence of food in the environment of the organism. Putting together our analysis of mistakes of misidentification and mimicry, and of the arbitrariness of markers, we can identify two kinds of correctness and mistake making in goal-directed systems. In the first kind, such as a response to M1, this response is only correct as long as M1 is associated with the environmental object or event for which action is appropriate. In the case of clotting this refers to M1 as a marker for injury. If M1 is associated with a different environmental object or event, the response is mistaken. In the case of clotting this refers, for example, to M1 associated with an atheromatous plaque. In the second kind, a response to a variation, M3, will be a mistake if it *impairs* reliable, accurate, and timely initiation. This mistake occurs not because injury is absent but because M3 is not an accurate marker for injury.[3]

Markers: Genetically Assigned and Acquired

If our aim to describe goal-directedness across biological systems is to be realized it must apply both to genetically directed processes, such as clotting, and acquired processes such as pressing a lever to acquire food. In turn, if markers are integral to goal-directedness, then they must operate in the same way irrespective of whether they are genetically assigned or are acquired. We claim not only that markers are found across the inherited–acquired demarcation, but that the properties we identified above with reference to the genetically assigned markers for platelet activation, make them uniquely suited to their acquisition as the result of experience.

In our earlier M1–M4 examples, the markers that we discussed are assumed to be established by genetic mechanisms. Inherited markers, and their receptors, have acquired their role in ensuring reliable, accurate, and timely goal-directed action through natural selection. M2s that are equally correct and M4s, which confer a clotting advantage, will also have been selected for. M3s to which a response is a mistake, will be selected out unless their operation under changed conditions confers selective advantage. These selective mechanisms occur, as we know, over generations.

Crucially, however, there is nothing in our analysis of markers that *requires* them to be established genetically. Markers with the same properties may be acquired. Furthermore, if markers can be acquired as the result of experience, they have the potential to be responsive to environmental changes or fluctuations that occur during the lifetime of the organism rather than solely from one generation to the next. Our previous discussion of M4 markers, ones that confer advantage, is especially relevant. A novel M4 which, let us now assume, indicates the presence of food over greater distances or time, in a way superior to M1 or M2 markers, may enhance the organism's goal-directed capabilities during its lifetime. In a changing environment, if M4 could be acquired as the result of experience, it would further enhance accuracy and anticipation of action to procure food.

Acquired associations of the kind that we envisage in this analysis of goal-directed behaviors are found widely in nature and have been extensively examined in a range of learning paradigms. For example, classical conditioning occurs when a neutral object is paired repeatedly with a reward such as food. This conditioned stimulus (CS) is an example of what we are referring to as M4, conferring advantage for goal-directed behavior in the environment. Consistent with our analysis, the CS can take many different physical forms and be processed through any of the senses. Tactile, visual, auditory, or olfactory stimuli of varying intensity each may become conditioned stimuli for the same reward. Once M4 has acquired the status of a CS, we can redraft the statement of goal-directedness based on Nagel, substituting M4 for food odor. This brings out that M4 must meet the same normative requirements as odor, i.e., it must enable the organism to identify and locate food accurately

---

[3] We are grateful to a referee for drawing our attention to a concern regarding Nagel's characterization of goal-directedness expressed by Matthen and Levy (1984). They argue that in a case such as autoimmune disease where the immune system attacks the cells of the body, it misidentifies self as foreign antigens, that the goal-directed mechanisms of the organism are activated, and hence Nagel "must say that the destruction of the self is the goal of this behaviour" (Matthen and Levy 1984). Our response is that nonself antigens are examples of markers for infectious agents, but if they are indistinguishable from the markers (antigens) on body cells, then the mistake of misidentification arises. In this instance destruction of the infectious agent remains the goal of immune behaviour, but it is misdirected because it responds to a marker for body cells that cannot be distinguished from a marker for an infection.

in the service of the goal of adequate nutrition for body tissues. Our Nagelian functional statement, highlighting the acquired M4 food marker and the normative qualifiers in bold, can be written as follows:

> **G$_{11}$**: a function of item *i* (***M4** receptor/sensor*) in system *S* (*organism feeding*) and environment *E* (*presence of food associated with **M4***) is *F* (***accurate** food identification and location*), presupposes that ***accurate and effective* action** ***(timely**, **well-directed** approach)* ***generated*** by *S (organism feeding*) is directed ***effectively*** to goal *G* (*ensuring **adequate** nutrition*) to the realization or maintenance of which *F* (***accurate** food identification and location*) contributes.

Recall that reward mechanisms are implicated in the way markers such as odor and taste facilitate evaluations of likely nutritional value, and these mechanisms in mammals rely substantially on neuronal dopaminergic activities. The reward function of these mechanisms does not require experience of the association between the marker and nutritional value of the food, and is assumed to arise from genetic mechanisms. Studies of the way neutral objects become markers of value for organisms—the CS—reveal that the neuronal mechanisms are very similar to those for the genetically established mechanisms. As we noted earlier (Schultz et al. 1997), in studies in which the activity of single neurons in the midbrain is recorded in alert monkeys, touching fruit or tasting fruit juice is followed by increased activity of dopaminergic neurons. The reward value of the nutritious object to the animal is registered as an increase in dopaminergic activity. The establishment of conditioned stimuli was investigated in experiments where monkeys had the opportunity to touch a lever after the appearance of a light. If that happened, initially by chance, they were rewarded with fruit juice (unconditioned stimulus). After repeated sequences of light, followed by lever pressing, followed by fruit juice, monkeys started to press the lever as soon as the light was illuminated, and without waiting for the juice. The light had become a conditioned stimulus. Furthermore, the illumination of the light now caused an increase in activity in the same dopaminergic neurons that were previously activated by sight or touch of food. Thus, a neutral, arbitrary object had become a marker for food, as a result of its acquiring reward value through dopaminergic neuronal activation (Schultz et al. 1997).

In the same way that genetically determined markers for survival goals will have consequences for selection over generations in the face of a changing environment, so also if an acquired marker results in the organism eating poison or moving towards a predator rather than food, the organism will not survive. Learning mechanisms that modify acquired markers allow for adaptation to a changing environment without threat to survival, and reversal of conditioned stimuli are well established (Stanton et al. 2021), again with the implication of dopaminergic mechanisms. When a conditioned stimulus is not followed by the predicted reward, neuronal dopaminergic activity is depressed, and when it is followed by more than the predicted reward it is elevated (Schultz 2016). Thus associations between markers and rewards may be acquired and reversed in response to changing environments, via dopaminergic neural processes.

Markers in Philosophical Context

Markers as Key to the Link Between the Organism and its Environment

Of the several possible philosophical contexts in which our proposal regarding markers could be placed, we briefly consider two here—the "field theory" of goal-directedness, and debates concerning representational content. The capacity of markers to be acquired is central to our understanding of the role of the environment in influencing goal-directed behavior. This can be seen by reference to the important "field theory" of Gunnar Babcock and Daniel McShea (Babcock and McShea 2021, 2023). They argue that the direction of goal-directed behavior is always outwards—towards the field which "is spatially larger than and envelopes the goal-directed entity" (Babcock and McShea 2023, p. 421). According to Babcock and McShea, there are many different kinds of field. We focus on what they term "ecological fields" comprising "food sources, climate conditions, predators, competitors, and parasites" (Babcock and McShea 2021, p. 8764); we do so because of their similarity to what we are referring to as the external environment of the organism.

Babcock and McShea refer to the field as "guiding" and as having "the capacity to direct" (Babcock and McShea 2021, pp. 8764–8765). Regarding the behavior of an organism, it is "its external field that tells it where to go" (2023, p. 421). In our account of goal-directedness, the link between goal-directed behavior and the field is specified by the marker and its connection to the field. It is not *directed* by the field. Consider an example that draws on their specification of ecological fields as containing food sources and predators. The ecological field of *C. elegans*, a worm that settles on rotting food, includes bacteria, some of which are pathogenic and others of which are food sources. According to field theory, the worms are directed by pathogenic bacteria to avoid them and are directed by food source bacteria to approach them. According to our proposal, by contrast, it is the markers for these organisms that account for the worm behavior, and as we have explained, markers can lead to misidentification and are open to revision, unlike the field objects themselves. For *C. elegans*, odor is a key marker either for predator or food source, but some pathogenic bacteria emit an odor that mimics that of a food source, leading to approach behavior and risk of infection. Thus pathogenic bacteria have "the capacity to direct" approach behavior or avoidance depending on correct or mistaken identification by the worm itself, based on the odor emitted. This is only possible because of the distinction between the object and the marker in the field. The implications of the odor marker can however change as the result of experience. Hermaphrodite worms that are initially attracted by their odor to pathogenic bacteria, learn from experience to show avoidance behavior to the odor (Peedikayil-Kurien et al. 2025). The key point, distinguishing significantly the mistakes approach from the field approach, is that the field object has not changed—it is the organism's *response* to the marker, and hence to the field, that has changed.

Markers and Representational Content

Although the focus of this article is on goal-directedness in general, across biological systems, our analysis makes contact with current discussions of mental representation and representational content. This is exemplified by the extensive philosophical discussion of "what the frog's eye tells the frog's brain," where the key question can be summarized as: does the optic nerve and the brain represent black dots in the visual field or food sources such as flies? (see Lettvin et al. 1959). According to our analysis of goal-directedness, whatever the frog responds to, be they black dots or other visual stimuli, are *markers* for food sources, with the properties that we have outlined. The link to content—the food source—depends on whether the marker provides the basis for effective action—approach to food. In a recent paper reviewing the representation-content debate centered on amphibian perception, Millikan (2024) summarizes the relationship thus: "the correspondence of a representation to what it represents must be a cause of the usefulness of the representation." The key words for our analysis are "correspondence" and "usefulness." Translated into our framework, "correspondence" captures the systematic association between the marker and the environmental object or event, allowing for the possibility of different markers corresponding to the same content. The "usefulness" of the marker depends on its *correctness* in leading to effective action in relation to what it represents.

Karen Neander's detailed account of how toads identify prey and predators brings out further similarities with our proposal (Neander 2017). She cites the work of Ewert (1987) using "dummy stimuli" to investigate what features of environmental objects toads use in responding either by approach, which is appropriate for food, or avoidance, which is necessary for a predator. He showed that toads respond to the size, shape, and motion of the stimulus. For example, they approach stimuli that are rectangles which move parallel to the longest axis. These experiments were conducted in the absence of any food or predatory sources, and the toads were responding to markers for those objects. Neander makes observations that are entirely consistent with our analysis, such as: "not all worm-like stimuli are worms, and not all worms are worm-like stimuli" (Neander 2017, p. 103). She does not, however, elaborate the normative properties of these stimuli, including that worm-like stimuli may lead the toad to approach mistakenly in the absence of food, or that a predator may mimic a food source by displaying the key features of a worm marker. Nor does she bring out, as we have, that the physical features of markers can vary substantially provided that they meet the relevant normative requirements for effective action—in this case approach to food and avoidance of predators. Further, Neander's analysis does not suggest a mechanism for acquired markers.

Intermediate, Internal Goals

All the examples that we considered earlier, of action directed towards survival goals, require there to be structures, including receptors, platelets, nerves, and limbs, already in place. These must be assembled by processes in which the goal is the completion of the structure—not only that, but the *correct* structure. The ultimate test of the structure's correctness is future action in the environment, but the intermediate goal is its correct assembly. This is important in two ways. First, it builds on the logic of the normativity of goal-directedness. If some goals have to be reached prior to their performing a survival function, such as stopping bleeding, providing adequate nutrition, or ensuring chick fledging, then the organisms must behave according to an *internal* standard of correctness. This is a standard that must be maintained *without* reference to whether it will succeed in underpinning action in the environment. Mistakes are departures from the standard, and if the organism is capable of identifying and correcting mistakes this also must be carried out relative to the standard. Secondly, the idea of an internal standard of correctness may provide us with a bridge to the many goal-directed behaviors of humans which are created, practiced, and enjoyed independently of survival goals.

Consider again the example of bird nest construction, not only because it is an exemplar of this phenomenon, but also because of its proximity to reproductive success. The standard of correctness which the nest has to meet is determined by its ability to provide a safe and benign environment for the hatching of eggs and the development of chicks prior to fledging (Healy et al. 2023). For this to be possible, nest construction must meet many normative requirements, such as that it is the right shape to contain the eggs, will generate the right temperature conditions for egg development, and that its location minimize the risk of predation. In the absence of feedback from whether a nest succeeds, which is not available during nest construction, the birds must make use of standards of construction, and if they correct mistakes then this must occur in compensation for departures from those standards. These are the intermediate goals that must be met if the further goal of reproductive success is to be achieved. Understanding of these phenomena is at the moment limited; however there is some evidence that in zebra finches (*Taeniopygia guttata*) males are the main bringers of materials, while their female mates modify the work of the males. We predict from our analysis that they do this based on a standard for nest building, but this remains to be investigated experimentally. In the case of bird nest construction, and more generally, the point is conceptual and not empirical. For a biological structure to be available for effective action in the environment it must have been assembled according to normative criteria (a standard of correctness) *prior* to the action. In each case the goal is the assembly of correct structures, and where mistakes are detected and corrected, this is relative to the standard of correctness for the goal.

The second kind of goal, also extensively studied in zebra finches—song learning—takes us closer to goals that are familiar to us as humans, such as the goals of learning to read and write, to dance, or perform a piece of music. We say this without prejudging or taking a position on the survival value of these human activities, and without disputing common ideas such as that music evolved to enhance social bonding (Savage et al. 2021). Our focus is on how to *characterize* the process, not its evolutionary origins. We also put to one side the potentially distracting question of whether conscious intentions are necessary in these human activities. Young male zebra finches learn their songs through a sensitive period by imitating their fathers, practicing their songs thousands of times, and refining them to create a unique song for courtship (Duffy et al. 2022).[4] During learning they produce fluctuating vocalizations. Based on recordings from dopaminergic neurons in the bird's brain, investigators have shown that vocalizations which more closely resembled the song learned from the father are accompanied by higher dopaminergic activity. This activity increases the closer the performance approaches the song of the father. Here the correctness of each song is evaluated against an internal standard, and improvements against that standard are rewarded via the dopaminergic system. Recall this is the same system that is implicated in establishing stimuli in relation to external rewards; only here the reward system is internal to the organism. The authors summarize the phenomenon thus: "spontaneous dopamine spiking can evaluate natural behavioral fluctuations unperturbed by experimental events such as cues or rewards" (Duffy et al. 2022, p. 1). They also make a broader point that is highly relevant to the possibility of a satisfactory conceptualization of goal-directedness: "Learning and producing skilled behavior requires an internal measure of performance" (Duffy et al. 2022, "In Brief"). The connections to the human activities mentioned above are tantalizing and eminently worthy of further investigation.

Conclusion

The phenomenon of goal-directedness, we have argued, needs to be characterized prior to proposing and examining possible explanatory mechanisms. Our central thesis is that the diverse phenomena that might be considered goal-directed do have a common property—their normativity. For many of them, normativity and action in the environment are inextricable. Systems and organisms must act when necessary in relation to environmental objects and events, and not at other times, not too early or too late, with actions that are well directed and effective in relation to goals such as adequate nutrition or avoidance of danger.

Our analysis of normativity leads to conclusions that, we suggest, further flesh out what is required for a productive account of goal-directedness. By identifying the indispensable role of markers for environmental objects or events in creating the conditions for reliable, accurate, timely, and effective actions, we open the space for an investigation of the commonalities between innate and acquired markers, and also of mechanisms for flexible, but also accurate, goal-directed actions in complex environments. We further exemplify the centrality of normativity to the possibility of goal-directedness by extending our analysis to the assembly of structures that are required for action in the environment and which must be completed to a standard ahead of time. Although we do not aim in this article to characterize human goal-directedness, and in particular the pursuit of goals for their own intrinsic value, we discuss one possible precursor to this phenomenon—the acquisition of bird song—in which

---

[4] The connection between zebra finch song learning and survival goals is a focus of much debate. There is good evidence that it serves territorial and mate selection functions, but this is by no means exhaustive and the possibility of further functions is widely discussed; see, e.g., Loning et al. (2023).

performance is evaluated against, and directed towards, an internal standard of correctness.

Acknowledgments This work was funded by the John Templeton Foundation within the "Agency, Directionality, Function" Program (grant no. 62220). The opinions expressed are solely those of the authors and not of the John Templeton Foundation.

Declarations

Conflict of Interest The authors do not have any interests to disclose that are directly or indirectly related to the work submitted for publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Babcock G, McShea DW (2021) An externalist teleology. Synthese 199:8755–8780

Babcock G, McShea DW (2023) Resolving teleology's false dilemma. Biol J Linn Soc 139(4):415–432

Bolton D, Hill J (2004) Mind, meaning and mental disorder: the nature of causal explanation in psychology and psychiatry. Oxford University Press, Oxford

Braithwaite RB (1953) Scientific explanation: a study of the function of theory, probability and law in science. Cambridge University Press, Cambridge

Chen Y, Lin YC, Kuo TW, Knight ZA (2015) Sensory detection of food rapidly modulates arcuate feeding circuits. Cell 160(5):829–841

Davies PS (2003) Norms of nature: naturalism and the nature of functions. MIT Press, Cambridge MA

Dresow M, Love AC (2023) Teleonomy: revisiting a proposed conceptual replacement for teleology. Biol Theory 18:101–113

Duffy A, Latimer KW, Goldberg JH, Fairhall AL, Gadagkar V (2022) Dopamine neurons evaluate natural fluctuations in performance quality. Cell Rep 38(13):110574

Ewert JP (1987) Neuroethology of releasing mechanisms: prey-catching in toads. Behav Brain Sci 10(3):337–368

Garson J (2019) What biological functions are and why they matter. Cambridge University Press, Cambridge

Gibbins JM (2004) Platelet adhesion signalling and the regulation of thrombus formation. J Cell Sci 117(16):3415–3425

Healy SD, Tello-Ramos MC, Hébert M (2023) Bird nest building: visions for the future. Philosophical Trans Royal Soc B 378(1884):20220157

Hill J, Oderberg DS, Gibbins JM, Bojak I (2022) Mistake-making: a theoretical framework for generating research questions in biology, with illustrative application to blood clotting. Q Rev Biol 97(1):1–13

Lee JG, McShea DW (2020) Operationalising goal directedness: an empirical route to advancing a philosophical discussion. Philosophy, Theory, and Practice in Biology 12:5

Lehtonen TK, Helanterä H, Solvi C, Wong BB, Loukola OJ (2023) The role of cognition in nesting. Philosophical Trans Royal Soc B 378(1884):20220142

Lettvin JY, Maturana HR, McCulloch WS, Pitts WH (1959) What the frog's eye tells the frog's brain. Proceedings of the IRE 47(11): 1940–1951

Loning H, Verkade L, Griffith SC, Naguib M (2023) The social role of song in wild zebra finches. Curr Biol 33(2):372–380

Matthen M, Levy E (1984) Teleology, error, and the human immune system. J Philos 81(7):351–372. https://doi.org/10.2307/2026291

Mayr E (1988) Toward a new philosophy of biology: observations of an evolutionist. Harvard University Press, Cambridge

Millikan RG (2024) Teleosemantics and the frogs. Mind Lang 39:52–60

Morales I, Berridge KC (2020) 'Liking' and 'wanting' in eating and food reward: brain mechanisms and clinical implications. Physiol Behav 227:113152

Morand-Ferron J (2017) Why learn? The adaptive value of associative learning in wild populations. Curr Opin Behav Sci 16:73–79

Nagel E (1979) Teleology revisited and other essays in the philosophy and history of science. Columbia University, New York

Neander K (1991) The teleological notion of 'function'. Australasian J Philos 69(4):454–468

Neander K (2017) A mark of the mental: in defense of informational teleosemantics. MIT Press, Cambridge

Nissen LA (1997) Teleological Language in the life sciences. Rowman & Littlefield, Lanham

Oderberg DS, Hill J, Austin C, Bojak I, Gibbins JM, Cinotti F (2023) Biological mistakes: what they are and what they mean for the experimental biologist. British Journal for the Philosophy of Science. https://www.journals.uchicago.edu/doi/pdf/10.1086/724444

Oderberg DS, Hill J, Austin C, Bojak I, Gibbins JM, Cinotti F (2024) Biological mistake theory and the question of function. Philos Sci. https://doi.org/10.1017/psa.2024.56

Okasha S (2018) Agents and goals in evolution. Oxford University Press, Oxford

Okasha S (2022) Goal attributions in biology: objective fact, anthropomorphic bias, or valuable heuristic? In: Corning PA, Kauffman SA, Noble D, Shapiro JA, Vane-Wright RI, Pross A (eds) Evolution 'on purpose': teleonomy in living systems. MIT Press, Cambridge, pp 237–256

Owyang C, Heldsinger A (2011) Vagal control of satiety and hormonal regulation of appetite. J Neurogastroenterol Motil 17:338–348

Peedikayil-Kurien S, Haque R, Gat A, Oren-Suissa M (2025) Modulation by NPY/NPF-like receptor underlies experience-dependent, sexually dimorphic learning. Nat Commun 16(1):662

Pittendrigh CS (1958) Adaptation, natural selection and behavior. In: Rose A, Simpson GG (eds) Behavior and evolution. Yale University Press, New Haven, pp 390–416

Rosenblueth A, Wiener N, Bigelow J (1943) Behavior, purpose and teleology. Philos Sci 10(1):18–24

Savage PE, Loui P, Tarr B, Schachner A, Glowacki L, Mithen S, Fitch WT (2021) Music as a coevolved system for social bonding. Behav Brain Sci 44:e59

Schlosser M (2015) Agency. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/agency/

Schultz W (2016) Dopamine reward prediction error coding. Dialog Clin Neurosci 18(1):23–32

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599

Stanton LA, Bridge ES, Huizinga J, Johnson SR, Young JK, Benson-Amram S (2021) Variation in reversal learning by three generalist mesocarnivores. Anim Cogn 24(3):555–568

Thomson JJ (2008) Normativity. Open Court, Chicago

Waddington CH ([1957]2014) The strategy of the genes. Routledge, Abingdon

Walsh PT, Hansell M, Borello WD, Healy SD (2013) Are elaborate bird nests built using simple rules? Avian Biol Res 6(2):157–162

Watts AG, Kanoski SE, Sanchez-Watts G, Langhans W (2022) The physiological control of eating: signals, neurons, and networks. Physiol Rev 102(2):689–813

Wedgwood R (2007) The nature of normativity. Oxford University Press, Oxford

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.