



Journal of Philosophical Investigations



University of Tabriz

Hegel's Internal Engine – Free Energy Minimization at Play in the Phenomenology of Spirit

Caius R. Gibeily 

Department of Biological and Biomedical Sciences and Department of Philosophy, Emory University, USA, Email: cgibeil@emory.edu

Article Info

Article type:

Research Article

Article history:

Received 20 July 2024

Received in revised form
01 August 2024

Accepted 03 August 2024

Published online 07
September 2024

Keywords:

Hegel, neuroscience,
predictive coding, free
energy principle,
phenomenology,
dialectics

ABSTRACT

This paper bridges contemporary neuroscience theories and Hegelian philosophy, centering on Karl Friston's Free Energy Principle (FEP). Neuroscience models like the Bayesian brain hypothesis and predictive coding depict the brain as a predictive machine, echoing Hermann von Helmholtz's concept of unconscious inference, where perception is shaped by prior knowledge. The FEP, rooted in information theory and statistical physics, suggests organisms minimize sensory surprise through unconscious and active inference, providing a model for behavior and explaining the purposiveness of biological systems. Some scholars assert that Georg W. F. Hegel's view of living beings in his Philosophy of Nature aligns with the FEP, portraying them as purposive and enactive systems. This paper extends this idea, proposing that Hegel's 'System of Science' in the Phenomenology of Spirit functions as a free energy-minimizing system. It discusses predictive coding and the FEP, establishing criteria for a system that minimizes free energy, and applies these criteria to Hegel's work. The paper argues that the dialectical narrative in the Phenomenology operates as a reflective system driven to minimize logical or conceptual free energy, ultimately advancing the spirit towards absolute spirit. This Hegelian predictive model generates expectations essential for dialectical progression.

Cite this article: Gibeily, C. R. (2024). Hegel's Internal Engine – Free Energy Minimization at Play in the Phenomenology of Spirit. *Journal of Philosophical Investigations*, 18(48), 81-94. <https://doi.org/10.22034/jpiut.2024.62559.3827>



© The Author(s).

<https://doi.org/10.22034/jpiut.2024.62559.3827>

Publisher: University of Tabriz.

Introduction

Many trending theories in neuroscience, including the Bayesian brain hypothesis, principles of efficient and predictive coding, and biased competition, revolve around the brain as a predictive machine that is engaged in continuous modeling of the world (Friston, 2010; Chalk et al., 2017; Pezzulo, 2021). An earlier form of predictive processing is found in the notion of unconscious inference proposed by Hermann von Helmholtz in which perceiving inevitably introduces perceptual priors that condition the perception of the object under scrutiny (Helmholtz, 1867; Wade, 2021). Karl Friston's free energy principle (FEP) is rooted in information theory and statistical physics and has found widespread traction in the biosciences (Friston, 2010). The principle integrates unconscious with active inference and offers a normative model of organism behaviour. The framework also offers an explanation for the teleonomy or purposiveness inherent to biological systems whose physicochemical and behavioral dynamics resist their destruction. Such systems achieve this goal by minimizing the surprise in their estimates of the sensory world (Friston, 2009).

Some contemporary scholars of Georg W.F. Hegel argue that his conception of living things in his *Philosophy of Nature* is one of living beings as intrinsically purposive and enactive natural systems (Boonstra and Slagter, 2019; Gambarotto, 2024). The *Phenomenology of Spirit* was Hegel's first major work in which he developed the introduction to his systematic speculative philosophy. This paper suggests that a form of free energy minimization is at play in the *Phenomenology*. The paper builds upon the only existing exploration of the FEP in Hegel's philosophy in Boonstra and Slagter (2019). The authors examine parallels to the FEP in Hegel's *Philosophy of Nature* where they present a compelling case for an FEP-like process at play in Hegel's understanding of animal organisms. The paper extends this perspective by arguing that Hegel's 'System of Science' (*System der Wissenschaft*) itself takes the form of a free energy-minimizing system.

The paper is divided into a brief discussion of predictive coding and the FEP, leading to a collection of minimum criteria that a system would need to meet to be considered a free energy-minimizing system. A case is then made for a possible reading of the FEP in Hegel's system, focusing on the Preface and Introduction of the *Phenomenology*.

It is argued that the dialectical narrative throughout the *Phenomenology* can be modeled as a type of reflective system immanently driven towards minimization of a form of logical or conceptual free energy. The inevitable result of this optimization in the *Phenomenology* drives the self-motion of spirit towards absolute spirit. A further consequence is that a Hegelian predictive model is also a generator of expectations, without which the dialectics could not progress.

Predictive coding and Karl Friston's Free Energy Principle

Neuronal predictive coding is the hypothesis that the brain predicts future sensory input rather than the pattern of sensory information. Activity in primary sensory areas reflects a mismatch signal whose amplitude is proportional to the disparity between the sensory evidence and the encoded sensory prediction, rather than a pattern of activity that signals the object's identity (de-Wit, 2010). Under a predictive processing framework, the brain actively anticipates and then compares the outcome to its priors, rather than being a passive recipient of incoming information from the sensorium. In hierarchical feedback circuits, there may be sequential feedback propagated from higher to lower levels that signal different types of prediction error (Rao and Ballard, 1999; Friston, 2009; Chalk et al., 2017).

Of course, without a method for updating their models, organisms would be unable to navigate their environments effectively and could have no certainty in their perception, given the volatility of the sensory world and inadequacy of out-of-date models to uncertain sensory information. The only information yielded would be a negative signal indicating 'x is not y', where x and y are prediction and observation, respectively. Instead of completely rejecting an inadequate model, leaving the perceiver without a starting point, the optimal strategy is a form of Bayesian inference that requires a starting point – the prior distribution – and calculates a posterior distribution, continuously updating the existing model. This overview gives some intuition as to why static predictions are unsuited in an uncertain world and, instead, the nervous system must continuously update its perceptual models to infer the true nature of the sensory world. Implementation of model update in nervous systems can involve modulation at multiple levels of the hierarchy, operating from changes in gene expression and protein regulation, modulating synaptic strength, up to reorganizations in brain network topologies and functional connectivity (Rao et al, 2024). Changes in synaptic weights given recent activity between pairs of neurons – in its simplest form, Hebbian plasticity – or large-scale adjustments in network activity through synaptic scaling are two such mechanisms that may transform prediction error into physiological changes in circuit dynamics.

Here, the FEP enters the scene as a principle for consolidating hypotheses on predictive coding, encompassing hierarchical Bayesian models of organism function (Friston, 2006, 2009 and 2010). While information theoretic free energy is a historical continuity of Helmholtz free energy in thermodynamics, it is distinct from thermodynamic Gibbs-Helmholtz free energy. The FEP states fundamentally that a self-organizing system will tend towards reducing its free energy. Informally, free energy reduction may be described as the efficient use of the information available in the dynamic system which reduces the amount of unused or wasted information that follows from encoding a surplus of potential information. Optimizing the free energy of the system involves setting a stable upper bound on entropy, or uncertainty in its predictions (Friston, 2006 & 2009). The statistical separation between the system's internal states and those of the surroundings is given by a Markov blanket. The FEP is formulated as having some set of n internal states, $\{\mu_1 \dots \mu_n\}$,

hidden states of the sensory environment, $\{\psi_i \dots \psi_m\}$, that explain the cause of the observed sensory distribution, and actions that follow from the internal state of the system, $a(\mu)$. The optimal state is the one that minimizes the free energy of its prediction given its sensory input (Friston, 2009) and past actions that follow from μ . Consequently, the self-organising system may opt to minimize its residuals through two compatible routes: changing its internal model to fit the available data or acting in such a way as to affirm its model predictions (i.e. changing the statistics of the sampled sensory space) (Friston, 2005 & 2006). The latter approach is consistent with enactivism whereby self-organizing systems are not just actively generating and testing their internal inferences against the sensory evidence but are also enactive, filtering their environment to confirm their prior beliefs. The enactive perspective includes brain, body and environment as a coupled system and provides a type of organismal teleonomy or purposiveness as a “self-organizing autonomous, autopoietic system” (Chemero, 2009, 152).

The Minimal Conditions for a Free Energy-Minimizing System

The FEP therefore provides a normative and unfalsifiable general model. Its formulation as a general principle is both germane in its applicability to a number of both biological and non-biological systems but also a limitation in the problem of discriminating between different putative free energy-minimizing systems. Both a thermostat and a human brain could be judged to perform free energy minimization. Distinguishing between FEP-compatible systems is beyond the scope of this paper; instead, a set of minimal criteria are proposed, which if satisfied, would point towards the system in question being a free-energy minimizing system *in nomine certe*:

1. Markov blanket – the system maintains a separation between its own internal states and the environment. The Markov blanket captures this separation in the statistics of the system’s internal states from its environment while still admitting the flux of information between system and environment (Friston, 2009; Pezzulo and Sims, 2021).
2. Mechanisms to reduce variational free energy – in addition to a separation of internal and external states, the system should have mechanisms for detecting and responding to the amount of variational free energy in the system by converging on a set of states with minimum free energy.
3. Generative modeling of expectation – the system models the expected distribution of the sensory evidence, termed the recognition density, which generates expectations about its current state (Friston, 2009 and 2010).
4. Active inference – the actions of the system should lead to a reduction in free energy. The resulting actions should either change the type of evidence the system is exposed to, such as through soliciting a particular type of sensory evidence or by changing its expectations (Friston, 2009, and 2009a; Pezzulo and Sims, 2021).

In the next section, Hegel's system and method in the Preface and Introduction are discussed, identifying elements of Hegel's system where it may be compatible with an FEP reading. It begins by building up from the prior examination of the FEP in Hegel's philosophy in Boonstra and Slagter (2019).

The Hegelian Free Energy-Minimizing System

The Free Energy Principle in Hegel's Understanding of Animal Organisms in the Philosophy of Nature

In their *Dialectics of Free Energy Minimisation*, Boonstra and Slagter (2019) propose a serious dialogue between Hegel's dialectics of the constitution of animal organisms and the FEP through a reading of plasticity in Catherine Malabou. They draw attention to the three moments in how organisms relate to the environment: “[r]eproduction passes through sensibility and irritability and absorbs them; it is thus derived, posited universality which, however, as self-producing is at the same time concrete singularity” (§353/*Philosophy of Nature*). The “contradictory organism” (Boonstra and Slagter, p. 10), arises in the negative relation of living things with their surroundings while maintaining an internal state in and for itself; organisms sublimate the two, thereby assimilating its surroundings.¹ As organisms reconstitute the experience of their organic and inorganic media, these moments are captured in habit (Malabou, 2000). These habits, or shapes, *Gestalt*, are taken to reflect internal states engendered by the FEP: “We thus find here an additional indication that both lines of thinking [Friston's states and Hegel's notion of habit] are closer than they may appear: both ‘state’ and ‘habit’ refer to the configuration of the anticipatory organism” (Boonstra and Slagter, 2019, 9). For the authors, the organism is by definition its anticipatory structure, potentiated through its form-receiving plasticity in making habit (Malabou, 2000). Upon encountering error that risks the organism's dialectical autodestruction, “the threat of the explosion of form structurally inhabits every form” (Malabou, 2008, quoted in Boonstra and Slagter, 2019), a mediated tension arises between the organism's formation and destruction of form that marks the dialectic progression, sublating the first two moments – form and the anticipatory error – to arrive at a new form. What is yet to be elucidated is whether the dialectics of free energy in organism determination in the *Philosophy of Nature* are reflected in the *Phenomenology*. The next subsection explores the properties of subject as living substance in the Preface, suggesting that a similar self-organisation is attested across different shapes of consciousness.

¹ “Now since the organism is directed towards the outer world as well as being inwardly in a state of tension towards it, we have the contradiction of a relationship in which the outer must be sublated. The organism must therefore posit what is external as subjective, appropriate it, and identify it with itself; and this is assimilation”

Hegel's System as Involving Self-Restoring Subject

Hegel's Preface, written after the completion of the remainder of the *Phenomenology*, offers important insight into his system and his goal of reconstituting a science of philosophy, as the "goal where it [philosophy] can lay aside the title of love of knowing and be actual knowing" §5. He warns that for the type of book the *Phenomenology* aims to be, it would be misleading to give a preface summarizing its results, given that laying out its result in advance without the stages preceding its actuality would be not to grasp that the results themselves are an end unintelligible without the means that got consciousness to those inferences. It is insufficient merely to have a collection of posited truths, since "[w]ithout this development, science has no general intelligibility [*Verständlichkeit*] and it seems to be the esoteric possession of only a few individuals" §13, marking an apparent contrast with the approach of his contemporaries. Hegel claims that "what is intelligible is what is already familiar and common both to science and to the consciousness alike" §13, thereby introducing the centrality of a cognizing consciousness into the science itself – the truths of the science must be immanent to consciousness itself in order to be intelligible, since intelligibility is by definition what can be grasped by consciousness (even if it does not appear so immediately).

In taking this stance, Hegel rejects the idea of intellectual intuitionism of a presupposed absolute, simply posited without development, and criticizes such views as he sees them in Schelling and Fichte (Pinkard, 2023; Bruno, 2023). He is also wary of what he perceives as the formulaic approach of the philosophical systems of his contemporaries and antecedents,¹ and seeks to escape the trap of "repetition of the same old formula" §15 that precludes the possibility of any movement towards an idea of absolute qua absolute; this may only proceed in following the "exposition of the system itself" §17 and to do so requires "grasping and expressing the true not just as substance but just as much as subject" §17. It is in the subject, the individual reader herself and consciousness that the truth of the whole is self-immanent.

Living substance is "the being that is in truth *subject*" §18 and it can only be "actual only insofar as it is the movement of self-positing, or, that it is the mediation of itself and its becoming-other-to-itself". Here, scholars have understood "actuality" (*Wirklichkeit*) as furnished by Hegel with a technical meaning of Aristotle's 'ἐνέργεια', the thing at work, real by expending energy (Pippin, 2022; Pinkard, 2023). To be constituted as subject, a living being must keep working towards its self-maintenance in the face of otherness. Unlike man-made devices, living beings are subject to their own internal action. Here, one finds a link to the first criterion of a free energy-minimizing system – subject qua living substance is to reject its own dissipation into its surroundings or

¹ However, if the way it spreads itself out is examined more closely, it turns out not to have come about as a result of one and the same thing giving itself diverse shapes but rather as a result of the shapeless repetition of one and the same thing which is only externally applied to diverse material and which contains only the tedious semblance of diversity §15.

becoming *merely* other and losing itself as in its death. It seems there is in effect a type of metaphysical Markov blanket here that divides the collection of processes constituting living substance as self-maintaining from what does not fall under the sheets. A further consequence is that living beings are a “self-restoring sameness, the reflective turn into itself in its otherness” §18. It should be recalled that similarly in the FEP, systems resist dispersion by tending towards as few unsurprising states as possible through (en)action and evidential inference – that is, they attempt to preserve their enduring self-similarity in the face of environmental perturbation.

Coincidentally, however, a living being cannot simply posit itself as itself without reference to otherness, a realization self-consciousness has before lordship and bondage.¹ The FEP interpretation is that living things also cannot deprive themselves of the external “sensuous world” for long, or, as Friston quips: “close [their] eyes or head for a dark room” (Friston, 2009a, p. 3), since organisms will eventually enter a surprising state like hunger. Part of their being is simply in the work they do to resist their own destruction. Hence, in the process of their self-maintenance, living beings are “pure negativity” within themselves (with hunger as a type of opposition to self-maintenance) and relate to objects negatively, as in Desire. While examples of sensory states are given here, this circularity in life’s teleonomy – “the circle that presupposes its end as its goal and has its end for its beginning, and which is actual only through this accomplishment and its end” §18 is perhaps a defining property of living substance generally for Hegel. It would be a teleonomy that is not simply in animal organisms and the mediations between them and the sensuous world but immanent in “[t]he life of God and divine cognition...as a game love plays with itself” §1 that is driven by “[t]he need to represent the absolute as *subject*” §23. The next section explores what form of free energy minimization may be at play across the successive shapes of spirit.

The Hegelian gradient of conceptual free-energy minimization through determinate negation

The paper has discussed that according to the FEP, the likeliest state (the recognition density) is chosen that reduces the variational free energy given the internal states of the system (Friston, 2009; Spezzulo, 2021). Under optimal Bayesian assumptions and plasticity in the system’s internal parameters, the result is a gradient of descent in free energy towards a minimum upper bound on surprise afforded by the available evidence and noisiness along the sensory channel (Isomura, 2023).

Turning to the *Phenomenology*, if we found such a gradient, one might ask what the optimization function that we as self-conscious would be attempting to solve by doing – and in the act of doing,

¹ “The sensuous world is thereby for it a stable existence, which is, however, only appearance, or is the difference which in-itself has no being. But this opposition between its appearance and its truth has only the truth for its essence, namely, the unity of self-consciousness with itself. This unity must become essential to self-consciousness, which is to say, self-consciousness is desire, full stop” §168. Self-consciousness has two moments – in the first, it takes the previous forms of consciousness from sense certainty and perception. In the second, it takes itself as object but discovers that to move beyond the empty assertion of ‘I=I’, self-consciousness must distinguish itself by opposing what is other to self-consciousness, namely, the sensuous world of appearances.

achieving – rigorous philosophical science? To make a case, therefore, for free energy minimization in the *Phenomenology*, it would not be concerned, at least exclusively, with empirical or quantitative free energy at play, for example, in biological systems interacting with their environment. Rather, it would take the form of conceptual or metaphysical free energy in logic involving “pure thought”. The result implies the development of an expanding system of non-contradictory philosophical truths, but not a reliquary of empirical or mathematical truths¹. What follows is the dialectical progression in which, for example, an apparent conceptual contradiction between the concept’s first moment and the next moment are sublated, thereby regarding both as necessary parts of the whole. Taking such an approach breaks with previous verificationist modes of thinking: “[i]t does not comprehend the diversity of philosophical systems as the progressive development of truth as much as it sees only contradiction in that diversity” §2.

To comprehend the progressive development while seeing each moment in the development of rigorous science is challenging for us being individuals who are accustomed to representational, “materialized thinking” §58 or, alternatively, in systems involving “non-actual thought”, such as formal logic. Hegel recognizes that breaking such patterns of thought “requires concentrated attention to the concept as such” §58. The implications are that for a system of rigorous science, self-consciousness should be aiming towards conceptual thinking, or that they are self-driven towards the goal of a form of conceptual thinking which in its wholeness is free of actual contradiction (even if parts of the whole present at first the semblance of contradiction). Consequently, in embarking on this route, it strives towards being a self-consistent system that is more than just a recapitulation of the ‘I = I’. The entire project insofar as it is a phenomenology of spirit involves a laborious gradient towards true knowing on the part of spirit where “immediate spirit” (§27, §679, §724) learns to know itself and to understand itself in its mediated role in world-history.²

The path brings natural consciousness to a bifurcation point in which it may fall back onto the truths it took as being certain. Such a route, one would suggest, is consistent with the enactive approach to minimizing the free energy by choosing to align itself to its prior expectations rather than update its truths. The other route involves updating its model by avoiding the trap of merely making a “return to the former truth” and moving forwards with difficulty to real truth in science. In a sense, the motivation to switch from the former to the latter may coincide with when sensuous consciousness begins to question its convictions at the start of the journey in Sense-Certainty on what it took as immediate sense certainty in the world of appearances but then discovered the

¹ To questions like, “When was Caesar born?”, “How many toise were there in a stadion and what did they amount to?”, etc., a neat and tidy answer is supposed to be given, just as it is likewise determinately true that the square of the hypotenuse is equal to the sum of the squares of the other two sides of a right-angled triangle. However, the nature of such a so-called truth is different from the nature of philosophical truths. §40

² In order to become genuine knowing, or, in order to beget the element of science which is its pure concept, immediate spirit must laboriously travel down a long path. §27

mediations between itself and the object and the internal mediations within both. These discoveries lead it on the spiral path of “despair” towards real truth, rather than the path of a stubborn or forgetful consciousness just reliving its past experiences without any development.

Along the arduous path, there is a certain teleological self-motion of the concept, elaborated more fully in the *Encyclopedia Logic*. Maybee suggests this movement for Hegel is really one driven by logical necessity, where the determinations themselves solicit their opposites and vice versa at the moment of understanding. It is a form of “natural deduction” that is self-propelled through rational motivation given prior experience (Butler, 2012): “the completeness of the forms of non-real consciousness will emerge through the very necessity of their progression and their interrelation” §79. Here, a contrast in the responses of the Kantian categories of cognition, with understanding (*Verstand*) and reason (*Vernunft*) is presented, where understanding handles the opposition by eliminating one of the comparands, whereas reason recognizes the necessity of determinate negation in grasping the whole, taking a dialectical stance (Pinkard, 2023). The final sentence of the section reveals that in determinate negation, “a new form has thereby immediately arisen” §79 due to the nothingness revealed from the previous form, which Pinkard interprets as the transition to a new form of life that replaces the earlier form but retains its experience (Pinkard, 2023). A point of parity between Hegel’s notion of determinate negation and prediction errors (in the broadest sense) is that both have some positive content in themselves but mark a negative relation to the first determination, stating what ‘is not’. One finds a similar argument raised by Boonstra and Slagter (2019) from Malabou, in which the first two moments of form and anticipatory error arrive at a new form of the organism.

The role of free energy in Hegel’s speculative philosophy could therefore be understood through the dialectical resolution of contradiction operating within and between the shapes of spirit. Within the predictive coding framework, determinate negations would constitute a type of error signal that are fed between layers. The logically necessary progression which Hegel takes simply to “come about on its own accord” §79 may be understood as a form of ‘maximum likelihood’ inference by consciousness. It takes the form that is least free energetic in terms of being logically most parsimonious (i.e. with the fewest extra presuppositions) while preserving itself in the next form given the experience consciousness has of its failures in the previous forms.

Free energy minimization at the level of cognition in the *Phenomenology* may both aid and be aided by a comparison with the psychology of cognitive dissonance. Cognitive dissonance arises when two apparently contradictory statements or beliefs are held simultaneously to be true, necessitating an update to the agent’s belief system (Festinger, 1957). To reduce dissonance, an individual may take the route akin to the understanding, negating one or the other of the beliefs through belief suppression or acting in a way that reduces tension between the beliefs, thereby bounding its surprise. Alternatively, it may involve a type of second-order meta-learning (Bateson, 1972). When faced with dissonance, people sometimes need to suppress their immediate error

response and instead engage in meta-learning, where the error is fed forward to a higher cognitive level, which in turn projects down to create a new belief model that resolves the dissonance (Kaaronen, 2018). In their predictive dissonance model, Kaaronen (2018) suggest that meta-learning prevents overfitting on limited information and recoding the original beliefs in a more generalized framework.

In summary, one has explored how spirit is on a necessarily ceaseless path towards self internal consistency in absolute spirit, capturing a form of teleonomy discussed above. The path, one has suggested, might be read as a path of gradient descent in conceptual free energy. One might object by raising the problem of equating teleonomy and teleology. The paper does not equate the two; rather, it suggests that the path has only the semblance of teleology in the context of spirit's development, just as pre-Darwinian theories of evolution, i.e. natural theology and Lamarckian evolution mistook organism ontogeny and phylogeny as directed towards some end design or utility. Instead, the semblance of teleology may arise from the work spirit does qua living substance towards its self-maintenance, i.e. through its teleonomy or intrinsic purposiveness. Such a reading may be attractive insofar as it does not presuppose an end-goal for spirit in history but allows it to emerge from spirit's immanent self-organization.

Determinate negation presents the semblance of contradiction reached by taking the current model to be true. It was argued to be characteristic of an error signal that signposts the failure of the current model, which, if consciousness avoids the trap of returning to old beliefs, is propagated forwards to a new form that reflects its past development. The new form also responds to the determinate negation that brought the destruction of the previous, unmodified form. The exposition given here points to a qualitative satisfaction of the second criterion of a free energy-minimizing system, one that follows a gradient of descent towards the smallest upper bound on its surprise. The role of expectation will be elaborated in the last subsection.

Expectation in Hegel's method of judging both knowing and the standard for knowing

The previous subsection explored the gradient of the system towards self-similarity and prediction error in the resolution of determinate negations that propel the system qua consciousness towards its goal of self-similarity. Still to be elaborated on is the role that prediction or expectation might play in Hegel's method for the *Phenomenology*. What contrasts the last two criteria with the second is the idea of 'active' rather than 'unconscious' or passive inference.

On the one hand, in unconscious inference (Helmholtz, 1867), the act of perceiving involves perceptual priors about the expected form of the object of perception, such that the object is inextricably bound by these inferences (Helmholtz, 1867; Wade, 2021). In vision, for example, these inferences may follow Wertheimer's Gestalt principles of perception, including the figure-ground distinction in which natural scenes are unconsciously deconstructed into two categories for the observer – agents in the foreground and static background (Wertheimer, 1923, Gluberman, 2017). These are implicit in perception and there is no sense that the agent seeks to test or self-

evidence its model of the world through action. While this paper certainly agrees with earlier arguments advancing a type of unconscious inference at play in a reflexive self in Sense-Certainty and Perception even before developed self-consciousness comes on the scene (Dahlstrom, 2013), one would, on the other hand, further propose that Hegel's method requires active inference. In Hegel, the active inference begins by acknowledging that the science of testing the truth of knowing requires an "investigation and testing of the reality of cognition" §81. The investigation entails testing the priors themselves in addition to the objects they purport to measure. Consciousness does not have the truth of knowing laid out before it, but must (we too as readers) actively determine this truth for itself. To do so however would be to invoke an established standard. Doing so is inadequate for Hegel as it implies a presupposition or an appeal to authority (Houlgate, 2012). Claims on *de facto* science also fail, since science is still in its early stages of development and thus yet to "justif[y] itself as the essence of as the in-itself" §81. When the object of examination is knowledge itself, what consciousness took as the in-itself (*a sich*) of knowledge was really *for* consciousness, being its object of study. Consciousness would then mistake the in-itself of knowing with what is actually *for* consciousness as its object. Here, Hegel makes two crucial points:

The essence or the standard would lie within us, and that which was supposed to be compared with the standard, and that about which a decision was supposed to be made on the basis of this comparison would not necessarily have to recognize [*anzuerkennen*] the standard. §84

The passage affirms that the standard is the one internal to consciousness itself and not an external imposition. Moreover, consciousness is not necessarily even cognizant of exercising a standard – a move reminiscent of unconscious inference. Indeed, the standard even deceives consciousness into thinking it knows what knowledge is in itself. By accepting the standard implicitly, consciousness falls into contradiction due to what it took as a true standard of knowing as it is in itself: "consciousness declares within itself to be the in-itself, or the true" §84. Hegel elaborates that consciousness "provides its own standard, and the investigation will thereby be a comparison of it with itself" §84, one that will be its own measure of knowing. The action of differentiating something like an object for consciousness from itself is a distinction immanent to the current form of consciousness. The distinction following from the model for making the distinction are those which consciousness must nevertheless investigate to determine if it has a true concept of knowing.

The method will consist, therefore, not just in examining knowing but also examining the very standards against which knowing is compared. In Hegel is the idea that each form of consciousness generates its own standard "for which the essence is something different from what was the essence for the preceding shape" §87. The standards always follow from but never directly precede the new form of consciousness. It is the standards that seem to change due to the emergence of a new form of consciousness. This progression was possibly reflected in the Preface where a

new form of consciousness is said to arise immediately from the earlier one. The standard, conversely, is a mediation of what knowing is taken to be for consciousness.

One would therefore draw from these observations that a standard is generated by consciousness that in turn reflects the experience it has of its earlier forms. The standard is thus an implementation of the current state of consciousness and what it takes to be the essence of true knowing. In this light, consciousness is a generative model of expectations that follow logically from the current form of consciousness; these expectations are the standards that are used to measure its knowing and in setting its determinations about something, for example, in distinguishing something from itself as an object. It follows, therefore, that consciousness could not begin to examine the truth of its knowing without also bringing along a self-imposed standard.

Given this framework and taking Hegel's standards as expectations for consciousness, it suggests a satisfaction of the final two criteria: a system that (3) generates expectation and (4) is engaged in active inference given the expectations it has generated of the world. Of course, the ideal position for consciousness would be to discard its standards or have a standard identical to the essence, but consciousness cannot completely discard them; philosophical consciousness instead updates them in a form that appears analogous to Bayesian inference with the goal of adjusting one's 'posterior distribution' to converge on the true essence of knowing.

The paper has explored a perspective on how Hegel's method for the system of science is one of making predictions, examining them and the objects measured against the predictions, and dialectically refining the underlying model of consciousness given the determinate negations that arise from the standard consciousness sets for itself.

Conclusion

If Hegel's system can be described normatively as a type of free energy-minimizing system, the next question would be how this reading impacts a reading of the main text of the *Phenomenology*. On one level, the FEP is poised to offer a unifying account of brain functioning, including perception, learning and action, and is well suited to understanding hierarchical dynamics that are intrinsic to biological systems. In the *Phenomenology*, such a reading could provide insight to a range of features in Hegel's system. Since Hegel's *Phenomenology* is deeply involved in charting the emergence of higher forms of consciousness as moments of spirit, the FEP is well suited to characterizing the self-organization of consciousness and how it maintains its coherency as it transitions through different forms of being. We saw as well how the apparent logical necessity between forms of consciousness driven by determinate negation might begin to be examined through a free energy reading. It would resemble the path a rationally motivated consciousness estimates would be optimal in terms of best resolving its prediction error qua contradiction, but in retrospect would not have been had it more experience.

Moreover, taking the free energetic approach provides an alternative exploration of the development of spirit towards its goal, namely, as the necessity of the system's minimization of

conceptual free energy across its development that strives to maintain itself in the face of determinate negations. In fact, one might even suggest spirit is what it is by virtue of its actualizing itself in its self-preserving work whose result leads towards absolute spirit. Future work could test the description of Hegel's system as a free energy-minimizing system through a more thorough reading of the core dialectical progressions in the main text, including in understanding the FEP in a historicist context, like the struggle for intersubjective recognition and historical ethical institutions.

References

- Bateson, G. (1972). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University of Chicago Press
- Bruno, G. A. (2023). From Time into Eternity: Schelling on Intellectual Intuition, pp. 1–15
- Butler, C. (2012). *The dialectical method: a treatise Hegel never wrote*. Humanity Books.
- Boonstra, E. A. & Slagter, H. A. (2019). The Dialectics of Free Energy Minimization, *Journal of frontiers in systems neuroscience*, 13 (42), 1–17. <https://doi.org/10.3389/fnsys.2019.00042>
- Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.
- Dahlstrom, D. (2013). The Self before Self-Consciousness: Hegel's Developmental Account. *Hegel Bulletin*, 34(2), 135–158. <https://doi.org/10.1017/hgl.2013.20>
- de-Wit, L., Machilsen, B., Putzeys, T. (2010). Predictive coding and the neural response to predictable stimuli. *J Neurosci*. 30(26), 8702-3. <https://doi.org/10.1523/JNEUROSCI.2248-10.2010>
- Festinger, L., Riecken, H. W. & Schachter, S. (1956). *When prophecy fails*. University of Minnesota Press.
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences*, 13(7), 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>
- Friston, K., Daunizeau, J. & Kiebel S. J. (2009). Reinforcement Learning or Active Inference? *PLOS ONE*, 4(7), e6421. <https://doi.org/10.1371/journal.pone.0006421>
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Friston, K., Kilner, J. & Harrison, L. (2006). A free energy principle for the brain. *Journal of physiology, Paris*, 100(1-3), 70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001>
- Gambarotto, A. & Mossio, M. (2024). Enactivism and the Hegelian Stance on Intrinsic Purposiveness. *Phenom Cogn Sci* 23, 155–177. <https://doi.org/10.1007/s11097-022-09823-9>
- Guberman, S. (2017). Gestalt Theory Rearranged: Back to Wertheimer. *Psychol.* 8, 1782. <https://doi.org/10.3389/fpsyg.2017.01782>
- Hegel, G. W. F. (2018). *The Phenomenology of Spirit*. Pinkard T. & Baur M., eds. Cambridge University Press.
- Helmholtz, H. L. (1948). Concerning the perceptions in general, first pub: 1867.
- Houlgate, S. (2012). *Hegel's Phenomenology of spirit: a reader's guide*. Bloomsbury Academic.
- Kaaronen R. O. (2018). A theory of predictive dissonance: Predictive processing presents a new take on cognitive dissonance. *Frontiers in Psychology* 9,1–15. <https://doi.org/10.3389/fpsyg.2018.02218>
- Malabou, C. (2000). *The Future of Hegel: Plasticity, Temporality, Dialectic* 15.

- Malabou, C. (2008). *What should we do with our brain?* Translated by S. Rand. Fordham University Press.
- Maybe, J. E. (2020). Hegel's Dialectics, *The Stanford Encyclopedia of Philosophy*, URL=<https://plato.stanford.edu/archives/win2020/entries/hegel-dialectics/>
- Pezzulo, G. & Sims, M. (2021). Modelling ourselves: what the free energy principle reveals about our implicit notions of representation, *Synthese*, Vol. 199 Springer Netherlands. <https://doi.org/10.1007/s11229-021-03140-5>
- Pinkard, T. P. (2023). *Hegel's Phenomenology of Spirit: A Guide*, OUP Usa.
- Pippin, R. (2022). The Many Modalities of Wirklichkeit in Hegel's Wissenschaft der Logik, in *Wirklichkeit. Beiträge zu einem Schlüsselbegriff der Hegelschen Philosophie*, pp. 43-58) <http://doi.org/10.5771/9783465106647-43>
- Wade, N. J. (2021). The vision of Helmholtz. *Journal of the History of the Neurosciences*, 30(4), 405–424. <http://dx.doi.org/10.1177/20416695211022374>
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* 39(1), 301–350. <https://doi.org/10.1007/BF00410640>

