Synthese (2025) 206:225
https://doi.org/10.1007/s11229-025-05311-0

ORIGINAL RESEARCH

Real patterns, the predictive mind, and the cognitive construction of the manifest image

Paweł Gładziejewski[1]

Received: 22 May 2025 / Accepted: 5 October 2025

Abstract

Dennett famously argued that constituents of the manifest (commonsense) image of the world are *real patterns*, where patternhood is grounded in data compressibility. This paper builds upon Dennett's original formulation by connecting it with recent work in computational cognitive (neuro)science. The aim is to use the notion of real patterns to shed light on the *genealogy* of the ontological commitments of the common sense, arguing that the processes by which humans learn and update internal models of the environment can be understood as extracting real patterns from sensory data. In particular, I trace a conceptual and mathematical progression linking Kolmogorov-Chaitin complexity and minimum description length to predictive coding, Bayesian inference, and predictive processing accounts of cognition. Then, I argue that this cognitive interpretation of Dennett's core idea suggests a structuralist (and Kantian) perspective on the relationship between mind and world, whereby the manifest image represents a structure present in sensory data. The paper concludes by sketching how this cognitive form of real-pattern view connects with (and possibly illuminates) metaphysical debates regarding the reality of two types of commonsense entities: selves and ordinary physical objects.

Keywords Dennett · Real patterns · Predictive processing · Self · Ordinary objects · Realism · Structural realism

✉ Paweł Gładziejewski
pawel.gla@umk.pl

1 Department of Cognitive Science, Nicolaus Copernicus University in Toruń, Gagarina 39, Toruń 87-100, Poland

Published online: 22 October 2025

1 Introduction

Dennett (1991) famously argued that the constituents of our commonsense, or manifest, image of the world are real patterns: regularities that objectively exist in virtue of mathematical facts about data compressibility. This paper aims to build on Dennett's original formulation of this idea by connecting it with recent work in computational cognitive (neuro)science regarding the cognitive mechanisms that underlie the construction of the manifest image itself. I will also sketch out some general philosophical conclusions from this cognitive interpretation of Dennett's patternism, which in some respects goes significantly beyond the original proposal.

Before I clarify my aims, let me recap the core proposal from *Real Patterns*. Dennett's (1991) original question was: what makes folk-psychological states *real* despite them not being identifiable with or reducible to lower-level (neural, computational, microphysical) states? His answer: beliefs and desires exist in the same sense that high-level patterns exist in Conway's Game of Life. The Game of Life is a cellular automaton in which cells on a grid turn on or off based on simple rules about their neighbors in the previous step. From those rules emerge stable, propagating configurations of "on" cells that behave like objects moving across the grid. According to Dennett, the reality of those high-level structures is objectively grounded in the facts about the compressibility of the information about the evolving grid. That is, as a matter of mathematical fact, one can describe and predict the evolution of the grid more concisely—using fewer bits—by tracking high-level objects instead of encoding the exact state of every cell at each step. Hence, these "non-fundamental" objects are real patterns in the Game of Life. Similarly, folk-psychological states like beliefs and desires exist as real patterns because they allow efficient, prediction-supporting compression of information regarding systems as complex as humans, without requiring the tracking of the processes (neural, computational, microphysical) that take place inside them.

The patternist framework can be naturally generalized beyond the domain of folk psychology to account for the metaphysical status of all kinds of non-fundamental entities. This includes entities that our manifest image is committed to, like, say, selves, tigers, or democracies (this is already present in Dennett, 1991; for broadly congruent recent work on ordinary objects, see Bird, 2023; Petersen, 2019). But perhaps the most fruitful applications of the core idea have been developed in the philosophy of science, where the notion of real patterns proved useful when accounting for inter-level or inter-theoretic relations *within* the scientific image (see Andersen, 2017; Burnston, 2017; Ladyman, & Ross, 2007; Millhouse, 2021; Seifer, 2023; Suñé, & Martínez, 2021; Wallace, 2012). In particular, the real-patterns framework began to act as a natural ally of projects that aim to establish or clarify a broadly non-reductive, weakly emergentist stance of how different "layers" of reality (or our conceptions of them) are metaphysically interconnected.

What I want to highlight is that both Dennett's original proposal and the work that it sparked tend to focus on the ontologies that we assume humans and scientific theories are already committed to. This approach takes these ontological commitments as given and then seeks to understand how they may fit together. My aim is to show how the concept of real patterns might elucidate the cognitive *genealogy* of such ontological commitments—an account of how they arise in the first place. In line with Dennett's original proposal, my focus here will be on the sort of entities that we represent in our conscious experience and/or in our intuitive theories, in short, the entities that populate the manifest image.[1] But I intend to argue that under theories that have recently become widespread in computational cognitive (neuro)science, the very process by which people *learn* and *update* their internal models of reality may be regarded as a form of *real pattern extraction*. For this, I will establish a natural mathematical and conceptual progression that links Kolmogorov-Chaitin complexity (which was the basis of Dennett's original proposal regarding the nature of patternhood) to the Minimum Description Length Principle, then *via* predictive coding and Bayesian inference, culminating in variational-free-energy-based accounts of cognitive architecture. Sect. 2 of this paper is devoted to this.

For Dennett, the real-patterns framework was a viable way to state and defend a moderate form of realism about folk psychological posits (and, by extension, about other domains to which the framework is applicable). I intend to build on this core idea by asking a question: what sort of realism, if any, does the present, cognitive interpretation of real patterns yield? I will argue that through extracting real patterns from sensory data, the brain could be understood as building up a structural description of its environment. That is, the internal models acquired through pattern extraction consist of structural claims about relationships between the inferred causes of sensory data. Interestingly, these structural claims can be regarded as *veridical* (at least within the realm of "phenomena") even in classical skeptical scenarios. This gives the view defended here a Kantian flavor and shows a natural way to bridge Dennett's original idea (at least under the present interpretation) to structuralism, as recently defended by Chalmers (2018, 2022). I develop this structuralist-Kantian spin on real patterns in Sect. 3.

Lastly, I plan to establish the present proposal as a tool for bringing additional clarity to certain debates that arise with regard to the metaphysical status of different domains of the manifest image. For this, I will outline two case studies. First, I will show how the cognitive interpretation of real patterns could elucidate the difference between realist and antirealist positions regarding the reality of *selves* (where debates persist despite different authors arguably agreeing on most underlying facts that could ground the (in)existence of selves). Second, I aim to show that the present proposal naturally dovetails with compression-based solutions to the "special composition" question that arises with respect to the status of *ordinary physical objects*. In this way, it plausibly vindicates a form of realism about ordinary objects, against mereological nihilism and the unrestricted composition view. The application of the present framework to specific metaphysical debates concerning selves and ordinary objects will be discussed in Sect. 4.

[1] In Sellars' original formulation (1962), the notion of "manifest image" referred to the overall framework through which humans ordinarily understand themselves and the world around them prior to the theoretical posits of science. Understood broadly, the manifest image includes not just the ontology of reality as it is consciously experienced or represented in (as we term them today) intuitive theories, but also the fact that humans conceptualize the world in normative terms related to epistemic reasons or moral and social rules. In the present paper, I set the issues of normativity aside and focus on ontology.

2 Predictive minds are pattern-extracting minds

We may start by taking a closer look at the construal of the notion of a pattern found in Dennett's (1991) seminal paper. Roughly, the story may be reconstructed to go like this. Patterns are patterns in something. That is, we first assume that there is a base-level substrate or a system in which a pattern could be discerned ("substrate" is the term I will use). Think of movements of individual molecules in a gas or the strings of words in texts used as a training corpus of a large language model (or individual cells in a cellular automaton grid mentioned above). Intuitively, for there to be a pattern, the components comprising the substrate need to exhibit some kind of regular, non-random structure. So, for example, "pressure" may be regarded as a pattern related to the average force exerted by molecular collisions on the container walls. Similarly, grammatical constructions (like noun phrases, verb conjugations, sentence structures) are patterns related to regularities in strings of words found in human writing.

Dennett then fleshes out this basic idea using the concept of information compression. For a substrate to exhibit a pattern, it needs to be possible to describe it more concisely than by describing the individual goings-on in the base components. Thus, characterizing a gas in terms of pressure is a way of compressing vast information about its microscopic states, like the velocities and momenta of individual molecules. A large language model capable of generating grammatically correct sentences is itself a heavily compressed version, implicitly embodied in model parameters, of its training corpus.

The next step is to bring clarity to the notion of compression in use. To be maximally abstract, let us assume that the pattern substrate is a string of binary digits. So, for a binary string of zeros and ones to exhibit no patterns means for it to be completely random, which in turn means it should be in principle impossible to compress it. Here, Dennett relies on the technical apparatus of algorithmic information theory, in particular on the concept of *Kolmogorov-Chaitin complexity* (often called "Kolmogorov complexity"; see Chaitin, 2006; Li, & Vitányi, 2008). So, a string is defined as incompressible—hence, patternless—if its K-C complexity is approximately equal to its own length. This means that the shortest possible computer program (written in a fixed universal programming language) that can generate that string is *as long as the string itself*. In other words, a string without patterns is impossible to effectively encode or redescribe in any other way than by repeating it verbatim. By the same token, what it takes for a string to exhibit a pattern is for its K-C complexity to be lower than the string itself. For example, a string "01010101010101010101" can be generated by a very short program: "Write '01' ten times." Of course, this abstract idea pertains to more concrete examples as well. Hence, characterizing a gas by pressure, the parameter set of a large language model capturing textual regularities, or describing the evolution of the Game of Life via high-level structures—they all serve as lower-complexity descriptions relative to their underlying substrates.

With this understanding of patternhood in place, Dennett claims that high-level ontologies can be vindicated insofar as they capture patterns through compression of lower-level (substrate) information. The reality of patterns, thus understood, is grounded in the objectivity of mathematical facts regarding K-C complexity, as it is not "up to the observer" whether, or to what degree, a program accurately describes a lower-level regularity. But it is also at this stage in Dennett's proposal (1991, pp. 33–37) that more observer-dependent, epistemic and pragmatic considerations are introduced. They are related to the aims and limitations of the systems or agents (like humans) that use high-level pattern descriptions. Consider that given a substrate, there will usually be multiple different ways to compress it efficiently. It is not the case that the shortest possible description will be the best when the goals and limitations of agents are taken into consideration. For example, these goals and limitations will decide the trade-offs between (1) how lossy a given compression is (how much lower-level information is lost under a given description), (2) how predictively accurate it is, and (3) how cognitively or computationally costly it is to use that compression for a given purpose. This part of Dennett's proposal highlights precisely the aspect that is my focus here: seeing high-level pattern descriptions as cognitive models that agents use when dealing with their environments. Here, I will set aside the which-patterns-are-best question (but see notes 5 and 6). Instead, I will focus on the main aim of the article and develop the idea that the compression-based notion of a pattern offers a viable story of how humans acquire their cognitive vantage point on the world. That is, I want to propose a way of thinking about the process by which people learn their commonsense ontologies as a process of extracting and refining high-level patterns.

To begin, as useful as the notion of K-C complexity is in fleshing out the very concept of a pattern, it must be noted that the K-C complexity of a string is not computable (Grünwald, 2007). Since cognitive processes, whatever they are, must be physically realizable and thus computable, K-C complexity cannot be directly computed, and thus used for learning or inference by the brain. To make real patterns more compatible with theories that could be acceptable in the context of cognitive science, we may instead opt to work with the quantity of Total Description Length (TDL). The crucial consideration is that TDL is conceptually close to K-C complexity—expressing the same basic idea—but it is also computable (or can be approximated) for specific model classes and coding schemes (Grünwald, 2007; Rissanen, 1989). TDL measures the combined number of bits needed to encode a model (M) for the data (D), denoted L(M), plus the number of bits needed to encode the data itself given that model, denoted L(D|M), such that:

$$TDL(M) = L(M) + L(D \mid M)$$

A lower TDL indicates a more compressed description, achieving a balance between the length of the model itself (L(M)) and its fit to data (L(DM)). To illustrate, consider the following 16-bit string: "0110111001101110." Now, take three candidate models that aim to efficiently capture the pattern in this string:

Model A: repeat "01" 8 times.

Model B: repeat "01101111" 2 times.

Model C: "0110111001101110".

To keep this heuristic example as convenient as possible, let us ignore the instruction overhead ("repeat *n* times") when calculating the length of each model. Under this simplification, the value of $L(M)$ is 2 bits for model A, 8 bits for model B and 16 bits for verbatim redescription of the original string that is model C.

Now, consider the term expressing the goodness of fit for each model, which quantifies how much additional information is required to accurately describe the string, assuming we know the model. Notice that the patterns described in models A and B *fail* to accurately encode some of the bits in the original string. For example, the string differs in ten positions from a string that would result from repeating "01" eight times, as described in model A. So, as simple as model A is, it imposes substantial additional cost in describing the data, hence $L(D\,|M_A)$ is relatively long. Again, to keep the example convenient, assume that the length of the description of the string, given a model, is equal to the number of bits that the model mischaracterizes when compared to the original string bit-by-bit (so, for example, it is 10 bits for model A). Now, model B makes only two such errors, so $L(D\,|M_B)$ is 2 bits. As a verbatim redescription, model C characterizes the original string completely and without errors, so $L(D\,|M_C)$ is 0. We can now calculate TDL for each model:

$$TDL\left(M_A\right) = L\left(M_A\right) + L(D|M_A) = 2 + 10 = 12\,\text{bits}$$

$$TDL\left(M_B\right) = L\left(M_B\right) + L(D|M_B) = 8 + 2 = 10\,\text{bits}$$

$$TDL\left(M_C\right) = L\left(M_C\right) + L(D|M_C) = 16 + 0 = 16\,\text{bits}$$

Importantly, TDL is tied to a model selection principle known as Minimum Description Length (MDL) principle (Grünwald, 2007). It simply states that when faced with competing models of data, one should choose the model that minimizes the TDL. In our toy example, MDL favors model B, as it balances model length and data fit better than its competitors.

As mentioned, what distinguishes the TDL/MDL approach from K-C complexity is its (at least approximate) tractability in practical contexts related to model evaluation and selection, for example in statistics or machine learning (Grünwald, 2007). This also allows us to bridge an abstract, mathematical construal of patterns-as-compressions to cognitive modeling. To hone in on this, let us start with considering the popular *predictive coding* approach as a specific instantiation of TDL/MDL. Predictive coding is first and foremost a family of compression algorithms (Spratling, 2017), and it entered the theoretical landscape of cognitive science as such (for classic work, see Rao, & Ballard, 1999; Srinivasan et al., 1982). The core logic that unites predictive coding algorithms is this: to encode a signal consisting of a series of data points, specify a prediction of that signal, and then encode only the errors between the actual data points and their predicted values. Consider again the simple example discussed above. We could interpret models A, B and C as defining rules for constructing a prediction of a binary sequence (for example, model A concisely specifies a 16-bit prediction: "0101010101010101"). To measure the error, each predicted binary digit could be compared to a corresponding digit in the actual sequence. For example, we could calculate the sum of the squared errors for each data point x_i (each bit in the sequence), given its corresponding model prediction (M_i). Summing these squared errors over all digits would return an overall measure of how well the model fits the data (for example, the squared sum of prediction errors for model A is 10). The assumption here: the larger the error, the longer its encoding. To express this more formally in the context of our toy example, the TDL of each model would be equal to a sum of the model length and the prediction error:

$$TDL\left(M\right) = L\left(M\right) + \sum_{i=1}^{N}\left(x_i - M_i\right)^2$$

Under such interpretation, applying the MDL principle would mean that we prefer encodings that keep the prediction rule as simple as possible (model length) while also minimizing the prediction error (data fit).

Notice that MDL can be interpreted as a rule for creating a compact description of data but also as a rule for selecting among candidate models of data. In this sense, MDL is as much about *inference* as it is about compression. By the same token, predictive coding has come to be recognized not just as a compression scheme but as an algorithm through which the brain could implement *Bayesian inference* (for a nuanced discussion, see Aitchison, & Lengyel, 2017).[2] Consider the formulation on which Bayesian inference is equivalent to maximizing the posterior probability of a model, given data, where:

$$P\left(M|D\right) \propto \ P\left(M\right)P\left(D|M\right)$$

This means that the probability of the model given the data (posterior, *P(M/D)*) is proportional to the prior probability of the model (*P(M)*) multiplied by the probability of observing the data given that model (*P(D/M)*). The claim, then, would be that the brain searches for an internal model that maximizes the posterior probability, given

[2] It needs to be noted now that in this paper, I proceed by assuming a particular interpretation of the predictive mind view (including predictive coding, Bayesian brain and predictive processing, variational inference, active inference). In particular, I assume here that (1) predictive theories discussed here aim at capturing the computational structure of actual neural mechanisms that underpin perception and cognition; (2) these theories construe the mind-world relation representationally, i.e. in terms of the brain building internal models of reality (see Gładziejewski, 2016, 2019). Importantly, there are influential approaches that interpret the predictive mind view in a way that rejects (1), (2), or (as is often the case) both at once. So, those positions regard the mathematical apparatus of the theory as useful modeling tool rather than a literal description of neural computational mechanisms, and/or they construe the theory in non-representational, 4E terms (see e.g. Bruineberg et al., 2018; Ramstead, Kirchoff, Friston 2020; von Es, 2020; see also Kirchoff, Kiverstein, Robertson 2022). This is not the place to argue in favor of my preferred approach, although I should note that a realist-representationalist interpretation allows to clearly cast pattern extraction as a computational process from which the manifest image literally emerges. Still, I think that even on instrumentalist/non-representational readings of the predictive mind view, one *can* meaningfully regard perception as pattern extraction, albeit in a weaker (perhaps diluted) sense. Hence, much of what is said in the present paper might arguably apply even on those interpretations. For example, one might claim that it is instrumentally useful (even if not literally true) to model the brain as a pattern extracting system, and to regard it as a system that relates to said patterns by being directly, non-representionally attuned to them in an enactive, embodied etc. manner (and thus implicitly committed to a certain ontology).

the data, by minimizing the prediction error. For example, in perception, minimizing the difference between internally generated predictions and actual sensory stimulation would serve as a mechanism through which the brain selects a model of the most likely causes of sensory stimulation (data). In this way, the core duality of compression and inference present in MDL is preserved in this context. I will return to the relationship between Bayesian inference and error minimization shortly, but first I want to focus on how Bayesian inference as such (regardless of whether it is implemented or realized by error minimization) connects to MDL.

A direct and rigorous connection transpires under the assumption that an optimal code is used, and thus the encoding length is directly related to probability or the information content, such that $L\left(x\right) \approx -\mathrm{log}P\left(x\right)$. In this way, the Bayesian prior and likelihood terms directly specify encoding lengths (for example, the higher the likelihood of data under the model, the shorter the description of data, given the model):

$$TDL\left(M\right) = L\left(M\right) + L(D|M) \approx \underbrace{-\mathrm{log}P\left(M\right)}_{\substack{\text{prior encoding}\\ \text{length}}} - \underbrace{\mathrm{log}P(D|M)}_{\substack{\text{like lihood encoding}\\ \text{length(data fit)}}}$$

Because the computational goal of Bayesian inference as construed (that is, maximizing posterior probability of a model) is equivalent[3] to minimizing $-logP\left(M\right) - logP\left(D|\,M\right)$, we can now see how Bayesian inference is equivalent to minimizing the TDL of a model, given data. That is, inferring the most probable model of data means finding the most efficient compression of that data. Bayesian inference is real-pattern extraction. Importantly, this core idea is preserved even if we remove the highly idealized assumption regarding the use of optimal codes[4]. After all, in line with MDL, the likelihood term in Bayesian inference already penalizes models that fail to fit the data. There are also aspects of Bayesian inference that implicitly penalize overly complex, or long-to-encode models. More complex models (for example, a double-peak mixture of Gaussians as opposed to a single-peak Gaussian, or a third-degree polynomial as opposed to a first-degree polynomial), while capable of fitting the observed data closely, are also flexible enough to fit a much wider range of other possible datasets. This high flexibility causes the posterior probability distribution for a complex model to become diffused across its large parameter space when conditioned on the specific observed data. Unless compensated for by an increase in data fit, this diffusion results in lower posterior confidence in any specific explanation, reflecting the model's tendency to overfit the data.

[3] This equivalence holds because maximizing the posterior probability *P(M | D)* is equivalent to maximizing the product *P(M)P(D | M)*. Due to the logarithm function being monotonically increasing, maximizing *P(M)P(D | M)* is equivalent to maximizing *log(P(M)P(D | M))* = log*P(M)* + logP*(D | M)*, which in turn is equivalent to minimizing its negative, *−logP(M) − logP(D | M)*.

[4] This assumption can break down in certain settings (for example, in formal logic and natural language, "*p* or *q*" is both longer to encode and conveys *less* information than "*p*") and it remains an open question whether the brain uses optimal or approximately optimal codes (however, see Barlow 1961; Benjamin, Zhang, Qiu et al. 2022).

It is at this point that all the threads introduced in this section can be tied together under the framework of predictive processing (PP), which is currently one of the most prominent approaches to modeling cognition and which was endorsed by Dennett in the later stages of his career (see Dennett, 2013; Dennett, 2017, Ch. 8). Here, I will focus on a mathematical formulation of PP that puts the concept of *variational free energy* (VFE) center stage (Friston & Kiebel, 2009; Parr et al., 2022).

The idea behind PP is that the brain *approximates* Bayesian inference by *minimizing* VFE. In particular, instead of directly calculating the posterior $P\left(M|D\right)$, the brain encodes *another* probability distribution, $Q\left(M\right)$, which is iteratively optimized to approximate the posterior. Formally, the point is to minimize the difference between those two distributions, formally measured as the Kullback-Leibler divergence between them (KL-divergence):

$$D_{KL}\left(Q\left(M\right) \parallel\ P\left(M|D\right)\right) = \sum\nolimits_{M} Q\left(M\right)\log\frac{Q\left(M\right)}{P\left(M|\,D\right)}$$

Notice, however, that the formula above assumes knowledge at the outset of the posterior $P\left(M|D\right)$. So the claim is *not* that the brain cannot reduce this KL-divergence directly. Instead, PP relies on the idea that the KL-divergence between the posterior and its approximation (*Q(M)*) is minimized *indirectly* through the minimization of VFE:

$$VFE = \sum\nolimits_{M} Q\left(M\right)\log\frac{Q\left(M\right)}{P\left(M,\ D\right)}$$

Here, the term $P(M,\ D)$—formally, a joint probability of data (sensory states) and their (postulated) worldly causes—expresses a *generative model* of the environment. PP rests on a postulate that a hierarchical model of this kind is encoded in feedback synaptic connections in the brain. The model hierarchy aims to capture the nested structure of causal dependencies between the parts of the environment and the way this structure generates the sensory input. The generative model and the approximate posterior $Q\left(M\right)$ are both assumed to be internally encoded and accessible for use in information-processing mechanisms of the brain.

According to PP, the generative model is used to predict the stream of sensory information. These predictions are propagated top-down and compared against actual input, giving rise to the error signal. Perceptual hypotheses (models) are readjusted to reduce the error. Minimizing the prediction error is a *mechanism* by which VFE is minimized. Hence, given that VFE minimization is a way to approximate the posterior, we may say that minimizing the prediction error is equivalent to searching for a $Q\left(M\right)$ that best approximates the true posterior under the generative model. *This* is why PP can be said to realize Bayesian inference (approximating the posterior) through a form of predictive coding (minimizing the prediction error). In other words, through minimizing prediction error, the hidden cause(s) of the sensory input are *inferred*. Over multiple iterations of this inferential process, the brain *learns* the (most likely) causal structure of the environment by setting the parameters of the generative model so that it effectively minimizes long-term, average prediction error.

The claim, then, is that the world as we represent it in perceptual experience is the result of this sort of inferential and learning process. This is the core cognitive mechanism through which the manifest image is constructed. Parts of this story have been used to explain the sense that the world is apprehended from a unitary first-person vantage point of "the self" (e.g. Hohwy, Michael 2017; Letheby, Gerrans 2017), and it is apprehended as flowing in time (e.g. Bogotá, Debbara 2023; Hohwy, Paton, Palmer 2016), consisting of spatially arranged (e.g. Gornet, Thomson 2024; Rorot 2021) and interacting ordinary objects (Gładziejewski 2023; Shwaninger 2022), with some of those objects also represented as subjects of folk-psychological mental states (e.g. Tamir, Thornton 2018; Veissière, Constant, Ramstead et al. 2020). Later on in this article, I will focus on two elements of this picture—the self and ordinary objects. For now, I want to finalize this more technical discussion.

What needs to be reiterated, then, is that PP, through combining predictive coding and Bayesian inference, inherits their status as a real-pattern-extracting process. Again, minimizing VFE *just is* a way of finding patterns in (sensory) data through information compression. This becomes apparent when we consider the following derivation of VFE (see Graves, 2011; Hinton, & Zemel, 1993):

$$VFE = \underbrace{\sum\nolimits_{M} Q\left(M\right)\log\frac{Q\left(M\right)}{P\left(M\right)}}_{\text{complexity}} + \underbrace{\left(-\sum\nolimits_{M} Q\left(M\right)\log P(D|M)\right)}_{\text{data fit}}$$

The complexity term here quantifies the cost incurred by the approximate posterior *Q(M)* for diverging from the prior distribution *P(M)*. It penalizes approximate posteriors that are "complex" in terms of how much they diverge from prior assumptions, effectively encouraging the posterior to stay close to the prior. That is, the complexity term here can be understood as quantifying the "length" or "descriptional complexity" of the revision from the prior (*P(M)*) to the new belief (*Q(M)*), thereby guiding the system towards updates that are as simple or "concise" as possible. Conceptually, this term is analogous to the model length term in TDL/MDL. The data fit term measures how poorly the model explains the observed data. That is, it encodes the average cost of explaining the data given the model, where the cost is higher if data fit is low. This corresponds conceptually to the "data length given the model" term in TDL/MDL. Thus, minimizing VFE balances the complexity of the explanation with its accuracy in fitting the data, mirroring the core idea behind MDL.[5]

[5] It should be noticed here how this point connects with Dennett's (1991) claim, mentioned earlier in the main text, that finding patterns in data has an inherently pragmatic element. That is, for a given raw substrate-level signal, there may be multiple possible ways to compress it that differ with respect to the trade-off they afford between simplicity and accuracy. The choice between them is, for Dennett, settled on pragmatic grounds related to one's particular goals or limitations. Interestingly, in predictive processing approaches, the free energy formalism *automatically* defines the preferred accuracy/simplicity trade-off. The very definition of variational free energy as an objective function, and the process of minimizing it, embodies this balance. However, the pragmatic element reappears when we realize that the specific architecture of the agent's generative model is not itself derived from first principles of free energy minimization alone. Rather, these foundational elements, which define the precise contours of the

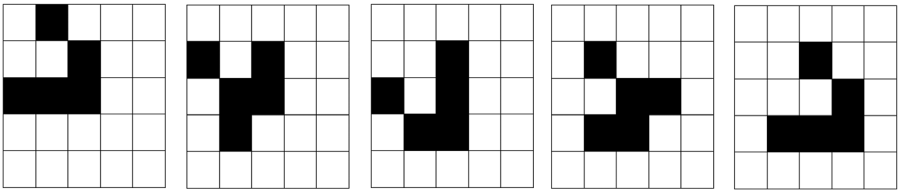

Fig. 1 Stages of a movement cycle of a glider in Conway's Game of Life

3 Patternism, structuralism, and realism about the manifest image

With the technical groundwork laid, it is useful to zoom out by drawing an analogy between the present proposal and Dennett's own classic Game of Life (GoL) example. In GoL, the raw substrate underlying the patterns is a grid of individual cells that turn on and off according to simple rules, based on the states of neighboring cells in the previous "generation." The higher-level patterns are objects that emerge from the rule-based shifts in the states of individual cells. For example, a "glider" is a small pattern of "on" cells that appears to move diagonally across the grid. It completes its cycle every four generations, returning to its original shape but shifting one cell diagonally (Fig. 1).

Now, in the PP framework, the incoming stream of sensory data (like retinal images or activations of the pressure-responsive cells in the skin) is like the grid in GoL. It is the base level: the raw substrate from which the brain extracts patterns. The brain's generative models compress these signals by minimizing prediction error, just as a description of a glider compresses the shifting states of individual cells. What we encounter in experience or in our intuitive theories—the manifest ontology of selves, tigers, and beliefs—consists of higher-level regularities discerned within the flux of sensory information. Constituents of the manifest image, then, are akin to gliders in the Game of Life.

Given that Dennett's patternism is usually regarded as a (moderate) form of realism with respect to domains to which it is applied, this may be taken to suggest that the present variant of patternism is also naturally read as advancing a kind of realism about the manifest image. Roughly, tigers and selves are real patterns in sensory data in the same sense that gliders are real patterns in the GoL grid. I think that this is the right direction, but when properly unpacked, the resulting view has philosophically important features that are not present—at least not explicitly or unequivocally so—in Dennett's own proposal.

A crucial consequence of the present view is that it advances a form of structuralism about how the manifest image is represented (see also Chalmers, 2018, 2022). Through building up generative models, the brain infers a model that captures recurring relationships and dependencies in the data: a description of a structure present in the data. Think of an analogy with gliders again. Gliders are elements of a high-level

free energy landscape the agent navigates, are themselves sculpted over evolutionary, developmental, or learning timescales by precisely the sort of pragmatic elements that Dennett highlighted.

structure that describes the substrate-level structure. For example, stating that a glider is present in a delineated part of the grid can be read as specifying the configurations of "on"/"off" cells one is likely to find in this part of the grid (namely, the configurations that correspond to distinct stages of gliders' movement pattern, as opposed to other possible configurations). Also, when a glider is located in a particular position on the grid, we may predict its diagonal "movement" towards future locations and thus implicitly predict probable cell configurations at later stages of the GoL evolution. So, gliderhood can be regarded as consisting, in part, in specifying the probability of substrate-level goings-on (this probabilistic spin on GoL may not be the most practical when one can simply run a simulation on a computer, but please keep in mind that I am merely using it as a useful analogy). But gliders also act as participants in a sort of causal structure that exists within the higher-level ontology, and through this implicitly compresses the substrate-level information. So, if I know that a glider has a blinker (a pattern of three neighboring "on" cells that oscillate by flipping orientation) on its path, I may predict that it will pop out of existence upon their "collision." Again, this is a high-level description that implicitly compresses substrate-level information. For example, the glider-blinker collision affords a prediction that a uniform group of off-cells will be present where a glider-congruent configuration of "on" cells would be expected otherwise.

Now, this intuitive example can be translated to how generative models represent the high-level structure present in sensory data. In a word, generative models specify probabilistic and causal structural descriptions of sensory patterns. Say that within my manifest image of reality, I perceive a door in front of me that I want to cross to enter another room. Now, just like gliders specify probable cell configurations, my perception of a door is a perceptual model or hypothesis that specifies predictions, spread over a multi-level processing hierarchy, about the probable outcomes within and across sensory modalities. That is, from the perspective of predictive processing, to be a "door" as humans perceive them is for such an object to bring about sensory feedback that is consistent with door-predictions (say, to appear rectangular under different vantage points, to emit knocking sound when knocked on, to be appropriately resistant to tactile pressure, etc.). But being a door also means participating in a higher-level structure of causal relations. This structure, in turn, implicitly predicts the unfolding of the substrate-level sensory input. For example, a door handle that can be acted upon by my hand in such a way that it results in opening the door, allowing me to cross—through engaging my body to bring about certain proprioceptive and kinesthetic effects—a previously obstructed location (arguably, the space itself may be considered as specifying a type of structure in this unfolding process; see also Chalmers, 2018; Schwitzgebel, 2019). Again, this tracking of high-level causal network of interactions structures the predictions about the unfolding of sensory information, which—if effective—manage to reduce prediction errors.

On this view, the constituents of the manifest image can be assessed as real in a way that aligns with Dennett's own patternist realism. The manifest ontology describes a structure present in sensory information, and this description arises from developing effective compressions of sensory information. These patterns, and the commonsense ontology based on them, are real in the sense that are not *freely* constructed. Here, I mean two things. First, the reality of patterns is constituted by objective, non-arbitrary mathematical facts about compressibility. Therefore, "manifest" structures, which arise from such compression, can themselves be regarded as non-arbitrarily or objectively *real patterns* within sensory information streams. The subject does not get to freely "decide": a pattern is objectively *not there* unless it is grounded in the structure of the signal.[6] Second, the patterns in the sensory signal are determined are the external causal source, and as such they constrain the generative model from outside (at least unless the model is used to run fully off-line simulations, as may be the case in imagery or dreams). In other words, the sensory patterns act as a stubborn and mind-independent factor that shapes the generative model (see Gładziejewski, 2021).

Further dissemination of this core idea reveals, however, features that are not present—at least not explicitly or obviously so—in Dennett's account of real patterns. Notice how the present proposal may be read as a certain constrained form of structural realism about the manifest image. That is, the generative model accurately depicts the *structural properties* of the process that generates the sensory patterns in the cognizers. Think, now, of a good old brain-in-a-vat scenario or any other skeptical scenario where the causal process that generates sensory input is systematically different from what we usually take it to be. By stipulation, in such scenarios, the sensory input itself matches exactly the non-skeptical scenario, making them epistemically indiscernible for the subject (at least based on perceptual evidence). Notice, now, that as long as the streams of sensory information (substrate-level) are matched between two kinds of scenarios, they will necessarily be equally compressible by equivalent higher-level (manifest) structural descriptions. That is, just like you, your twin-brain-in-a-vat effectively compresses sensory streams through prediction error minimization, using a generative model that represents commonsense entities. Hence, real patterns present in a non-skeptical scenario are equally present in an equivalent skeptical scenario. If so, the door that your envatted twin opens to enter another room in her virtual world is as real as the door you open to enter another room. This amounts to a variant of veridicalism about skeptical scenarios, developed in recent years by Chalmers (2018, 2022). Insofar as the manifest ontology is a structural description of sensory patterns, this description is equally veridical in skeptical and non-skeptical scenarios, provided the sensory inputs themselves are identical in both.

Notice that in this approach, the assessment of "reality" of the entities that furnish the manifest image is settled by investigating the relation between the substrate-level sensory patterns and higher-level structural description of those patterns. In this sense, the statements about reality (or veridicality), are grounded in relations between cognitive states that are internal to the epistemic perspective of the cognizer—that is, relations between what happens at the sensory periphery and the internal structural models induced from this (although we need to keep in mind that the former is itself constrained by an external source). In a broadly Kantian framework, this sort of realism is confined to *phenomenal* realm: the world as it appears to the cognizer.

[6] This is not to deny that discerning and tracking patterns in sensory data also has a pragmatic aspect to it (as Dennett himself admitted). That is, out of a larger space of in-princple mathamatically possible compressions of sensory signal, human subjects—with their goals and limitations—converge on those compressions that are the most useful and computationally tractable (see also notes 5 and 9). In this sense, the manifest image is clearly "a view from somewhere." However, I do not think that conceding this takes away from the overall realist stance developed here.

Whatever external reality (that is, the external source of the signal) is generating sensory streams, the cognitive agent can only know or accurately represent the *structural* facts about the *effects* that this reality is producing in her.[7] Beyond this type of epistemic reach lies the *noumenal* world of things as they are irrespective of how they appear from the epistemic vantage point of the subject (see also Schwitzgebel, 2019).

To make this more concrete, consider three types of noumenal facts that escape the manifest image in this way. First, there are facts about *realizers* that fill the structural roles described by the generative model (Chalmers, 2018). For example, the door-role (something that affords passage, has certain visual/tactile properties, interacts causally in specific ways, etc.) is presumably realized by a physical object in one's actual case, but is realized by a data structure on a supercomputer for one's brain-in-a-vat twin. One cannot discern what realizes the structural description of sensory data just through compressing such data. Second, there may be purely intrinsic, non-structural properties of the world that would, by definition, be beyond the grasp of any structural model. Third, there are presumably structural facts about the underlying reality from which the sensory input is ultimately sampled. For example, elements of the *scientific* image of the world (say, the quantum wave function of the universe) might count as such. From the perspective of the cognitive agent constructing the manifest image, these deeper structures could be loosely considered "noumenal" in the limited sense that they are not directly induced from sensory data and represented in the manifest image. Importantly, however, these deeper structures would fail to be noumenal in the strict Kantian sense that places things in themselves beyond the scope of knowledge, including scientific knowledge.[8]

It is not my intention here to dwell too much on speculations about the world *an sich*. What is important is the overall architecture of the realism that, I think, naturally follows from the cognitive interpretation of patternism presented here. In line with the overarching philosophical intentions behind Dennett's original introduction of real patterns, what emerges is a form of realism about the manifest image, one with a distinctly structural-Kantian face. It *is* realism, even if not "industrial-strength."

[7] As pointed out by an anonymous reviewer, the account of cognitive construction of the manifest image presented here is developed from an individualistic and classically empiricist perspective, leaving out the notion that knowledge arises from *collective* and cultural processes. However, this individualistic stance stems from the particular aim and scope of the present paper, rather that commitments regarding the nature of rationality or general epistemological considerations. My focus here is on the part of the manifest image that arises from early-developing, relatively low-level cognitive capacities—like object perception and minimal self-modeling (see Sect. 4)—that emerge largely independently of social or cultural scaffolding. These capacities, while they may precede and condition our entry into the communal "space of reasons," do not themselves rely on participation in rational social discourse. (The narrative self, also discussed in Sect. 4, may be a exception to this, insofar as the capacity for constructing self-narratives may arguably be heavily socially mediated).

[8] In this context, it also ought to be noticed that the present proposal remains compatible with naturalized metaphysics or certain structuralist-realist threads in philosophy of science (e.g. Ladyman, & Ross, 2007). In particular, it leaves open the possibility of *vindicating* the manifest image through the scientific image by establishing how the high-level "manifest" structural description of reality may be grounded in structural description(s) of reality provided by scientific theories (perhaps with a special emphasis on fundamental physics). If we assume realism about scientific theories on some independent grounds, this sort of move could additionally strengthen the realist stance on the manifest image. Here, I set this aside as a separate project from the one I am pursuing.

The manifest image is, on this view, a structural description of reality, grounded in sensory data. It is not up to us that the sensory data we obtain embody a certain non-random, patterned structure. The contents of our internal (generative) models are, in turn, constrained by this sensory structure; that is, the way sensory data shape the structure of generative models is itself non-arbitrary. The patterns in the signal determine, by way of generating prediction error to be minimized, which higher-level or coarse-grained descriptions effectively compress it. In this sense, the manifest ontology is not freely conjured up or imagined. The manifest image is an objectively constrained image (see also Gładziejewski, 2021).[9]

4 Manifest entities as patterns: the self and ordinary objects

The aim of this section is to make the preceding discussion more concrete by applying cognitive patternism to two entities that feature in the manifest image: selves and ordinary physical objects. In particular, the goal is (1) to explain how the representations of such entities are constructed in the cognitive system and thus what it means for the manifest image to be committed to their existence; (2) to show how this sort of view fits into and potentially illuminates existing metaphysical debates about the relevant domains of discourse. Importantly, the discussion to follow is meant as an initial sketch that establishes the fruitfulness of this approach rather than provides mature proposals—a prolegomenon to a future cognitive-structuralist metaphysics of the manifest image, if you will.

4.1 The self as a pattern

On the predictive processing view discussed in Sect. 2, the brain constructs a structural model of the causes of sensory input by extracting recurring patterns from the sensory signal itself. Crucially, some of these recurring features are *endogenous*—they originate from the prediction-error-minimizing system itself. As I type these words, the tactile feedback from my fingertips arises from my pressing the keyboard; my visual input depends in part on self-initiated eye movements; the interoceptive signals indicating hunger arise from within the same organism that engages in prediction-error minimization. These immediate, self-related sensory patterns are, in turn, embedded within longer-term, endogenous regularities: the writing-related tactile and visual inputs I receive today ultimately originate from my self-conception as

[9] Importantly, this still leaves room for still more fine-grained ways of qualifying the elements of the manifest image as real (or unreal). Dennett himself (2013) noted that predictive processing allows us to regard some of the elements of the manifest image as "projections" of idiosyncratic, observer-relative properties onto external reality. Perhaps the elements of the manifest image lie on a continuum between more descriptive and more projective patterns in sensory signals, such that the projective end consists of entities or properties that are also more mind-dependent (in a sense that would require additional work to fully flesh out). For example, when perceiving a baby, the way I represent its *spatial location* is closer to the descriptive end, while my perception of its *cuteness* is more a projection of my own (predictions about the) emotional reactions to the baby's presence (see also Wiese, 2025).

an academic, just as my interoceptive signals reflect my enduring habit of skipping breakfast.

On the predictive view, then, the self is inferred and represented as a common, single cause that underlies certain (endogenously generated) regularities in the sensory signal (Hohwy & Michael, 2017; Letheby & Gerrans, 2017). That is, the internal generative model aims to efficiently predict the complex flow of sensory input by distilling it into a compact representation of "I" or "me." This self-model achieves compression by positing a single, stable entity as the common cause of many endogenous patterns; by doing so, it makes future self-related actions and interoceptive states more predictable, thereby reducing the information needed to encode the ongoing stream of self-generated sensory experience.

This self-modeling process takes place across a hierarchical structure that tracks regularities unfolding across different time scales. In this sense, it can be said that there are multiple self-representations, corresponding to sensory patterns that can be discerned at different temporal horizons, and hence are represented at different levels of the model hierarchy (see also Gallagher, 2013 for a broadly similar view). The simplest way to understand this point is by distinguishing the minimal self from the narrative self (Hohwy & Michael, 2017; Letheby & Gerrans, 2017). The minimal self emerges from the predictive modeling of fast-timescale sensory dynamics and is experienced as the sense of being an embodied agent located in space, as well as the owner of ongoing mental states. The narrative self arises from the brain's integration of endogenous regularities over longer timescales, modeling the organism as a temporally extended agent with goals, memories, habits, personality traits and social roles. It underpins the experience of personal identity across one's individual life history.

So, in light of the patternism on offer here, the commitment to selves in the manifest image is, fundamentally, a commitment to the existence of real patterns within the sensory signal that are best compressed by a self-model. But are such selves-as-patterns *real*? Are they real *enough*? That is, is having such a self sufficient for having a *real self*? What becomes relevant here are meta-level assumptions regarding how stringent requirements are imposed on entities such as selves to count as real. Traditionally, these requirements are quite inflationary. In both Western and Eastern philosophical lineages, to be a self is to be a persisting, unified, substantial entity—a core of identity, a soul, an ātman. This kind of self decidedly appears to be more than just a pattern. So, to the extent that we require the self to be a substance (or substance-like in some robust enough sense) to count as real, the self-as-pattern view is best interpreted as a kind of anti-realism about selves (see Letheby & Gerrans, 2017, who draw this exact conclusion). In Buddhist terms, what emerges may be thought of as a type of no-self view.

To sharpen this point, think of the self-as-substance as an entity that is thought to ground one's numerical identity over time, such that me at time t_1 and me at t_{1+n} are numerically identical if and only if we are the same self-as-substance. On the patternist view, however, it's unclear where such a substantial entity could naturally fit. Take the following analogy with the Game of Life, inspired by Derek Parfit's work on personal identity (Parfit, 1984, Ch. 11–12). Suppose two gliders, A and B, are moving across the grid. We pause the game and exchange a portion—say, two—of their constituent black cells (see Fig. 1). Upon resumption, the game proceeds exactly as it would have without the intervention. The two glider patterns, in Dennett's sense of structures that support effective compression, persist. Insofar as gliders are *patterns*, the true description of the situation still holds: there is a glider at one location and another glider at the other. But are these post-transplantation patterns *numerically identical* to their pre-transplantation counterparts? Is one of them still numerically identical to glider A, and the other one to glider B? There appears to be no further fact within the GoL "world" that could non-arbitrarily settle this question—no fact about the evolving grid itself, the higher-level structures, nor any fact determined by those. For a pattern in GoL, persistence over time is not underwritten by some deeper entity beneath the shifting cell configurations.

Parfit, drawing on similar thought experiments, argued that personal identity over time does not depend on some "further fact" about the persistence of a substantial self (Parfit, 1984, Ch. 11–12). Just as there is no deeper fact of the matter about which glider is numerically which after the cell-swap, there is no deeper self determining whether a future person is numerically identical to a past one. Instead, what exists are various degrees of psychological and physical continuity and connectedness. From the present point of view, we could say that the brain uses these continuities—conceived as regularities unfolding over time, much like gliders—to construct its representation of the self. That is, the unfolding constellations of sensory signals exhibit regularities that allow for their compression via a self-model. For this process to be effective, the relevant pattern must, as a matter of fact, reside in the signal. In this sense, the self is real as a structural fact about the effects that reality is producing at the sensory boundary of the cognitive system (a view consistent with how Hohwy & Michael, 2017 interpret the predictive processing story; see also Beni, 2019 who defends a different but related version of non-substantivist realism about selves, grounded in structural realism in philosophy of science). Under these less traditional, deflated criteria, such a self is as real as any real pattern.[10] Importantly, we can now see how this overall story can be regarded as anti-realist with respect to selves-as-substances (in line with Letheby, & Gerrans, 2017) and realist with respect to selves-as-patterns (in line with Hohwy, & Michael, 2017). The conflict between those positions is apparent, and arises at a meta-level discussion about which criteria of reality are more reasonable to use for this particular domain. Tradition suggests substantial selves, but it seems like the patterns-are-enough view has been gaining traction over recent decades.

[10] Notice how this view dovetails with some other threads of Dennett's work on the nature of self. First, it arguably allows the self to be (at least partially) self-constituted, in that sensory signals that support the self-representation are a sort of self-fulfilling prophecy (see the concept of selves as self-writing novels in Dennett, 1992). For example, the cognitive system makes some high-level prior assumptions about a narrative self and then, through active inference, acts in a way that conforms to those assumptions, effectively confirming the model (see Hohwy, & Michael, 2017). Second, we can think of cases where the usual coherence or continuity of self-related patterns breaks down. For example, in certain dissociative states, the long-term behavioral patterns may become disjointed such that they are better compressed by model that posits two or more narrative selves residing within one body (in line with the discussion in Dennett, 1992).

4.2 Ordinary objects as patterns

The manifest world is also furnished by ordinary physical objects located in time and space, like cupboards, cats, pizzas, and doors. We primarily know, or at least seemingly know, that the world contains such entities from perception.

What I want to focus on is the question of the reality of such ordinary objects insofar as it is informed by considerations regarding mereological composition. Ordinary objects are composed of parts. A cat is mereologically composed of paws, whiskers, eyes, a tail, internal organs, etc. (Beyond what is discernible within the manifest image, it is presumably further composed of cells, proteins, molecules, subatomic particles, etc.). But by what principle does a collection of objects compose a larger, distinct object? Suppose that I find myself in a room that (apparently) contains a cupboard, a cat, and a pizza. Why assume that this is the way of parceling the world into objects as they really are? Why not think that a veridical description of what is going on contains an exotic (from the human point of view) object that is composed of cat whiskers, an outer crust on the pizza, and all dust particles coating the surface of the cupboard? The manifest image gives primacy to one unique way of carving up the world. The question is: is this carving at the joints, revealing genuine divisions in the world as it really is? There are two ways to deny that this is the case. According to *mereological nihilism*, composition never occurs—there are no composite objects, neither cats, nor pizzas, nor any exotic composites. According to *mereological universalism*, any arbitrary collection of objects composes a larger object. We would be wrong to think that ordinary objects enjoy any privileged metaphysical status. The commitment to the existence of cats becomes trivialized, for contrary to appearances, cats exist in exactly the same way as the "whiskers + crust + dust" amalgam, or any other such exotic entity.

A proponent of a revisionist position regarding ordinary objects (like mereological nihilism or universalism) must face the challenge of explaining why perceptual evidence seems to favor a unique way of parsing the world into entities. But a relatively straightforward move for meeting this challenge seems available. One might claim that, contrary to what may seem at first, we have no perceptual evidence that uniquely favors an ordinary object ontology. That is, if instead of ordinary objects, the world contained no composite objects (only atoms arranged object-wise) or contained all possible mereological compositions as objects, our perceptual experiences and perceptual beliefs would be exactly the same. Hence, perception provides no grounds for favoring an ordinary object ontology. This way, any claimed perceptual justification for an ordinary object ontology becomes debunked (see, for example, Korman, 2014).

In recent years, a novel solution to the composition problem has emerged, one that aims to provide a principled ground for favoring the commonsense ontology of objects over revisionist alternatives. Of crucial relevance here is that this new proposal puts the concepts of compression and real patterns center stage (Bird, 2023; Petersen, 2019; for a particular variant that makes use of the apparatus of the Free Energy Principle, see also Beni, 2025). According to one formulation of this approach, "the Xs compose an object Y when the Xs are a non-divisible maximal collection of objects such that relevant information about Y compresses corresponding information about the Xs in an efficient albeit possibly lossy way" (Bird, 2023, p. 686). Essentially, the "maximal" condition here means the collection of Xs is not merely an arbitrary part of a larger system that contains the same compressible pattern, ensuring the object has distinct boundaries. The "non-divisible" condition signifies that this collection forms an integrated whole for the purpose of compression, one that cannot be divided without loss of efficiency in compression. Space precludes a detailed discussion of this view, but suffice it to say that it promises to provide principled solutions to certain problematic cases regarding objecthood, avoiding many of the issues found in previously advanced accounts (Bird, 2023; Petersen, 2019).

To see how this works in practice, think of different ways of carving up the world into discrete objects as different ways of compressing a vast amount of information about what is going on. The parts of a cat (organs, limbs, cells) exhibit kinetic, functional, and biological correlations. Stating that "the cat is sleeping" or "the cat is a predator" compresses information about the state and coordinated activities of relevant parts. The cat is a real pattern. Now think about an ontology containing the "whiskers + crust + dust" composite object. There is little patterned behavior unifying those items, and hence no significant information compression achieved by grouping them. The parts are not kinetically correlated in a shared way beyond coincidence, nor functionally integrated, nor do they share a common fate that would be simplified by positing "whiskers + crust + dust." The "cost" of defining and tracking such disarranged amalgam would be high, with minimal payoff in terms of efficiency of compression. Knowing something about "whiskers + crust + dust" as a whole tells us little new or efficiently compressed information about the properties of the whiskers *as whiskers*, the dust *as dust on a cupboard*, or the crust *as part of a pizza*.

Here, the story above naturally coincides with the cognitive patternism developed in this paper. In particular, from the predictive mind perspective, the compression-based solution to the composition problem can be naturally reinterpreted to account for how *perceptual mechanisms* parse sensory signals into *representations* of ordinary objects (Gładziejewski, 2023; Schwaninger, 2022). For example, when perceiving a cat, the brain infers a "cat-object" as a common cause that explains correlations between signals that contain information about lower-level sensory features (like color, shape, movement, etc.). Tracking such an object is a way of picking out invariants—patterns—in the signal, creating a simple, efficient description that groups them under a single entity. Within the predictive/variational framework, this grouping occurs because a generative model that posits a single, coherent object achieves a better balance of accuracy (predicting the correlated features) and complexity (a single cause) than alternative models, thereby minimizing overall prediction error. In other words, the compression-based *metaphysical* principle described above determines facts about composition, such that ordinary objects enjoy a privileged metaphysical status. But this principle can also be regarded as capturing the core *computational* rule by which those very objects are inferred and represented in perception.

From this point of view, we can also explain how perception can be regarded as evidentially sensitive to whether the world contains ordinary objects (see also Gładziejewski, 2023 for an in-depth discussion). That is, the ordinary, cat-containing ontology and more mereologically exotic ontologies (like the one containing "whiskers + crust + dust") can be treated as competing generative models. The point, then, is that the models representing ordinary objects are better at extracting *real patterns* from sensory signals. For example, a "cat" model provides a superior compression of the sensory data, which means that a perceptual system operating on principles of prediction error minimization will converge on models representing "cats" rather than "whiskers + crust + dust" amalgam. The bundle of sensory features consistently associated with a "cat" exhibits strong internal correlations and co-predictability; for instance, the movement of its paws is highly predictive of the trajectory of its torso and tail, forming a unit where parts mutually inform the whole. No such cohesion is tracked by the "whiskers + crust + dust" model. By its very nature as an arbitrary collection, this latter model presents no unified pattern for efficient compression; its disparate components lack the regularities that would allow a model to achieve any significant reduction in complexity regarding the sensory signal. Furthermore, the ordinary "cat" model tracks a cohesive bundle of features that has a predictive boundary. That is, adding unrelated sensory data (like that from a nearby pizza crust) wouldn't reliably enhance, and could disrupt, the predictive effectiveness of the "cat" pattern.

5 Conclusions

Dennett's original goal behind the idea of real patterns was to articulate a viable moderate realism for non-fundamental entities, particularly those that feature in the manifest image of the world. Here, I have tried to extend that project by (1) connecting it to a computational story about the cognitive mechanisms that generate the manifest image; (2) arguing that the resulting view naturally can be interpreted as a form of structural realism (supplemented by certain Kantian themes); and (3) applying this to two specific examples of manifest entities: selves and ordinary physical objects. Before closing, a caveat. Although I have developed the present view in the context of Bayesian, predictive processing, and variational models, the very idea that cognition is a form of information compression is not new or unique to those models (for example, see Barlow, 1961; Wolff, 2019). So the present view may generalize beyond the particular approaches that I have focused on.

Acknowledgements I thank the two anonymous reviewers for their constructive and insightful comments on the earlier version of this paper.

Author contributions The author is the sole contributor to the paper.

Funding No funding was received to assist with the preparation of this manuscript.

Data availability Not applicable.

Declarations

Competing interests No competing interests to be reported.

References

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227.

Andersen, H. K. (2017). Patterns, information, and causation. *Journal of Philosophy*, *14*(11), 592–622.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). The MIT Press.

Beni, M. D. (2019). *Structuring the self*. Springer.

Beni, M. D. (2025). Constraining the compression: Thermodynamic depth and composition. *The Philosophical Quarterly*, *75*(2), 396–406.

Benjamin, A. S., Zhang, L. Q., Qiu, C., Stocker, A. A., & Kording, K. P. (2022). Efficient neural codes naturally emerge through gradient descent learning. *Nature Communications*, *13*(1), 7972.

Bird, A. (2023). Restricted composition is information compression. *Philosophical Quarterly*, *73*(3), 677–700.

Bogotá, J. D., & Debbara, Z. (2023). Time-consciousness in computational phenomenology: A Temporal analysis of active inference. *Neuroscience of Consciousness*, *2023*(1), 1–12.

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese*, *195*(6), 2417–2444.

Burnston, D. C. (2017). Real patterns in biological explanation. *Philosophy of Science*, *84*(5), 879–891.

Chaitin, G. (2006). *Meta Math! The quest for Omega*. Knopf Doubleday Publishing Group.

Chalmers, D. J. (2018). Structuralism as a response to skepticism. *Journal of Philosophy*, *115*(12), 625–660.

Chalmers, D. J. (2022). *Reality+: Virtual worlds and the problems of philosophy*. W. W. Norton & Company.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy*, *88*(1), 27–51.

Dennett, D. C. (1992). The self as a center of narrative gravity. In F. Kessel, P. Cole, & D. Johnson (Eds.), *Self and consciousness: Multiple perspectives* (pp. 103–115). Erlbaum.

Dennett, D. C. (2013). Expecting ourselves to expect: The bayesian brain as a projector. *Behavioral and Brain Sciences*, *36*(3), 209–210.

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of Minds*. W. W. Norton & Company.

Friston, K. J., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B*, *364*, 1211–1221.

Gallagher, S. (2013). A pattern theory of self. *Frontiers in Human Neuroscience*, *7*, 1–7.

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, *193*, 559–582.

Gładziejewski, P. (2019). Mechanistic unity of the predictive Mind. *Theory & Psychology*, *29*, 657–675.

Gładziejewski, P. (2021). Perceptual justification in the bayesian brain: A foundherentist account. *Synthese*, *199*, 11397–11421.

Gładziejewski, P. (2023). Un-debunking ordinary objects with the help of predictive processing. *British Journal for the Philosophy of Science*, *74*(4), 1047–1068.

Gornet, J., & Thomson, M. (2024). Automated construction of cognitive maps with visual predictive coding. *Nature Machine Intelligence*, *6*, 820–833.

Graves, A. (2011). Practical variational inference for neural networks. *Neural Information Processing Systems (NIPS) 2011*, 2348–2356.

Grünwald, P. D. (2007). *The minimum description length principle*. The MIT Press.

Hinton, G. E., & Zemel, R. S. (1993). Autoencoders, minimum description length and Helmholtz free energy. *Neural Information Processing Systems (NIPS)* (pp. 3–10).

Hohwy, J., & Michael, J. (2017). Why should any body have a self? In F. de Vignemont, & A. Alsmith (Eds.), *The subject's matter: Self-consciousness and the body* (pp. 363–392). The MIT Press.
Hohwy, J., Paton, B., & Palmer, C. (2016). Distrusting the present. *Phenomenology and the Cognitive Sciences*, *15*(3), 315–335.
Kirchhoff, M. D., Kiverstein, J., & Robertson, I. (2022). The literalist fallacy and the free energy principle: Model-building, scientific realism, and instrumentalism. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/720861
Korman, D. Z. (2014). Debunking perceptual beliefs about ordinary objects. *Philosophers' Imprint*, *14*(13), 1–21.
Ladyman, J., & Ross, D. (2007). *Everything must go: Metaphysics naturalized*. Oxford University Press.
Letheby, C., & Gerrans, P. (2017). Self unbound: Ego dissolution in psychedelic experience. *Neuroscience of Consciousness*, *2017*(1), 1–11.
Li, M., & Vitányi, P. (2008). *An introduction to Kolmogorov complexity and its applications*. Springer.
Millhouse, T. (2021). Really real patterns. *Australasian Journal of Philosophy*, *100*(4), 664–678.
Parfit, D. (1984). *Reasons and persons*. Clarendon.
Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*. The MIT Press.
Petersen, S. (2019). Composition as pattern. *Philosophical Studies*, *176*(5), 1119–1139.
Ramstead, M. J. D., Kirchhoff, M. D., & Friston, K. J. (2020). A Tale of two densities: Active inference is enactive inference. *Adaptive Behavior*, *28*(4), 225–239.
Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
Rissanen, J. (1989). *Stochastic complexity in statistical inquiry*. World Scientific.
Rorot, W. (2021). Explaining Spatial purport of perception: A predictive processing approach. *Synthese*, *198*, 9739–9762.
Schwaninger, A. C. (2022). Predicting ordinary objects into the world. *Philosophical Psychology*, *37*(8), 2134–2157.
Schwitzgebel, E. (2019). Kant Meets cyberpunk. *Disputatio*, *11*, 411–435.
Seifer, V. A. (2023). The chemical bond is a real pattern. *Philosophy of Science*, *90*(2), 269–287.
Sellars, W. F. (1962). Philosophy and the scientific image of man. In R. Colodny (Ed.), *Frontiers of science and philosophy* (pp. 35–78). Pittsburgh University.
Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition*, *112*, 92–97.
Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London, B*, *216*(1205), 427–459.
Suñé, A., & Martínez, M. (2021). Real patterns and indispensability. *Synthese*, *198*, 4315–4330.
Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social Mind. *Trends in Cognitive Sciences*, *22*(3), 201–212.
van Es, T. (2020). Living models or life modeled? On the use of models in the free energy principle. *Adaptive Behavior*, *29*(3), 315–329.
Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, *43*, e90.
Wallace, D. (2012). *The emergent multiverse*. Oxford University Press.
Wiese, W. (2025). Conscious perception as augmented reality. *Social Epistemology*. https://doi.org/10.1080/02691728.2025.2496747
Wolff, J. G. (2019). Information compression as a unifying principle in human learning, perception, and cognition. *Complexity*, 2019, 1879746.