Predictive Hermeneutics

Graham Johnson and Tom Rutten

The first step in the direction of truth is to understand the frame and scope of the intellect itself, to comprehend the act itself of comprehension. [...] The first step in the direction of beauty is to understand the frame and scope of imagination, to comprehend the act itself of esthetic comprehension.

—Joyce, *A Portrait of the Artist as a Young Man*

If I am to listen to your esthetic philosophy give me at least another cigarette.

—ibid.

ABSTRACT

Predictive processing, a meta-theory from cognitive science and computational neuroscience, proposes the mind as a hierarchical system which predicts reality as it unfolds (Clark 2016, Hohwy 2013, Friston 2006). This prediction occurs at levels ranging from the "merely" sensory to the highly conceptual. Here, we extend this thesis in order to understand the hermeneutic process as it relates to textual and artistic encounters. We argue that one of the foundational functions of contemporary artistic and literary production is to inform and exploit the mind's predictive system. We further show how this conceptualization of the art encounter as a cognitively rich interaction between the artwork and mind, coupled with the predictive processing framework of cognition, helps explain a host of traditional literary, aesthetic, and art historical values, including ambiguity, defamiliarization, and reversal.

TABLE OF CONTENTS

§1 Into the predictive mind

1.1 Historical precedents

> The schematism by which our understanding deals with the phenomenal world... is a skill so deeply hidden in the human soul that we shall hardly guess the secret track that Nature here employs.
>
> — Kant, *Critique of Pure Reason*

We can think of the inferential or "predictive" mind as an elaborate feedback system of induction and deduction: pulling patterns out of data and interpreting data based on previously pulled patterns (i.e. acquired knowledge). Though the problem of inference is "as old as recorded Western thought," (Griffiths, Kemp, and Tenenbaum, n.d.) we attempt here to sketch out a handful of precedents, of particular relevance to this paper, for the cognitive-scientific predictive mind, taken from philosophy, psychology, hermeneutics, and aesthetics.

Both Kant and Piaget worked prominently with the term *schema*, referring to a pattern of thought that organizes information (categorically, relationally, etc.) into a mental structure or "interpretive framework" that influences attention and learning (Swanson 2016; Beni 2017). The schema anticipates predictive processing models by arguing that an organism's understanding of reality requires a reconciliation of top-down models with bottom-up data.[1] In *Driven By Compression Progress* (2009), Schmidhuber notes a forerunner of his "cognitive compression" work (an alternate formulation of predictive coding[2]) in Piaget's theories of explorative learning and assimilation, where "new inputs [are] embedded in old schemas" (Schmidhuber 2009). Psychologist Alison Gopnik has also taken up this frame, arguing that children's cognitive development occurs in a manner strikingly similar to shifts in scientific theory, and can be understood as the accumulation—in an approximately Bayesian manner—of informal theories about the world as a structured causal system (Gopnik 1996).

Psychological set, and its sub-concept perceptual set, originated in the 1950s and 60s as a description of how situational interpretations can be highly influenced by expectations, and how perception is best understood as an "active process involving selection, inference, and interpretation" (Hill 2001). The perceiver has expectations and allocates attention deliberately (referred to as the *selector* function of perception). He also knows how to classify, understand, and name selected data, as well as draw inferences from them (*interpreter* function). Factors influencing perceptual set include expectations, emotions, motivation, and culture. One of the more famous experiments on perceptual set involves the numbers *12-13-14* written vertically in criss-cross with the horizontally written figures *A-13-C*. readers exposed to the vertical figures interpreted the middle figure as the numeral "13"; readers exposed to the horizontal array interpreted the middle figure as the letter "B" (Bruner and Leigh Minturn 1955). Perceptual set was anticipated by the Gestalt psychology of the early 20th century, which argued that the mind actively organizes incoming perceptions, through their interrelation, into gestalt wholes.[3]

Around the same time, the emerging field of cybernetics had begun "framing purposive behavior" in systems of all kinds as "governed by feedback," the seeking out of novelty with which to fuel said feedback system (Turner and Larson 2015). In literary theory meanwhile, hermeneutic work was undergone by thinkers like Hans-Georg Gadamer, Wolfgang Iser, and Ernst Gombrich, work which helps inform our own schematic theorizing of art encounters. As noted by L. Kesner, Gadamer's hermeneutical scenario highlights the "ongoing and dynamic" quality of a textual encounter, whereby parts inform whole informs part and the whole history of the reader is activated in the exchange (Kesner 2014). Gombrich (1960), meanwhile, emphasized the role of the schema in both the production and reception of visual works.

[1] Kant's philosophical work may anticipate predictive processing in ways beyond the concept of top-down modeling. Swanson (2016) notes that the concepts of hyperpriors (see §5.6) and generative models, among others, also stem from the German philosopher. Fazelpour and Thompson (2015) even go so far as referring to cognitive science's predictive mind as a "Kantian brain." Additionally, there is some speculation of an indirect influence *by* Kant on the development of predictive processing theories, specifically via Hermann von Helmholtz's inference machine. For discussion, see (Swanson 2016; Beni 2017)) and (Swanson 2016; Beni 2017).

[2] See Schmidhuber (2009) Appendix A1, "Predictors vs. Compressors"

[3] Parallels to Gestalt theory are one of many bridges (alongside similarities between machine learning processes and theories of predictive coding in human minds) to Peli Grietzer's "Theory of Vibe," which analogizes modernist art to the representational process of a trained autoencoder (Grietzer 2017).

In §1.2, we explain the fundamentals of predictive coding, prediction error minimization (PEM), and its speculated extension via concepts of Bayesian inference. Section 1.3, which focuses on the interfacing of PEM and hierarchical Bayesian models, will be especially of pertinence to those with research interest in high level predictive processing and the Bayesian brain. Non-technical readers, primarily interested predictive processing as a lens into art interactions, may benefit from skipping to §2. A glossary section at the conclusion of the paper provides a references for Bayesian and predictive terms which may be unfamiliar.

1.2 From predictive coding to error minimization

Predictive models of perception originated in computational neuroscience in the late 90s a way of explaining low-level sensory processing such as in the visual and auditory fields (Rao and Ballard 1999). However, cognitive philosopher Andy Clark and neuroscientist Karl Friston have proposed that predictive processing and error minimization explains cognition at all levels of the brain and, through Friston's free energy principle, constitute a fundamental property of life (Clark 2013; K. J. Friston 2013; Clark 2017).

Here, we proceed from several stipulations of these theories: one, that the predictive dynamics theorized of lower-level sensory processing also operate at higher levels of cognition such as conceptual learning, and two, that the inference systems central to lower-level predictive coding are similarly in play at higher levels of cognition. Justification for this extension of predictive coding from the merely sensory into the conceptual can be found in (Hohwy 2013; Clark 2013; K. J. Friston 2013; Clark 2017)), under the term *predictive error minimization*, Clark ((Hohwy 2013; Clark 2013; K. J. Friston 2013; Clark 2017), (Hohwy 2013; Clark 2013; K. J. Friston 2013; Clark 2017)) under the handle *predictive processing*, and (Hohwy 2013; Clark 2013; K. J. Friston 2013; Clark 2017)) as *free energy*. As noted in (Clark 2016)), it has also been dubbed *active inference* in theories emphasizing its extension to motor control,[4] and *hierarchical layered coding* by those emphasizing the organizational hierarchies believed to structure the different levels of prediction work.

In this paper, we'll use the term prediction error minimization (PEM) as an umbrella term for cognitive-scientific models (such as Hohwy, Clark, and Friston's) postulating that the brain actively actively uses predictive structures, organized hierarchically from low-level sensory to high-level conceptual domains, to anticipate future events. Moreover, PEM theorizes that, through a self-evaluation of its own anticipatory response, the mind builds accurate models of the world by reducing the amount of error in its predictions of sense data. For our purposes, predictive error minimization is congruous with Clark's conception of hierarchical predictive processing.

We can understand PEM as a "systematic bridge linking three of our most promising tools for understanding mind and reason: cognitive neuroscience, computational modelling, and probabilistic Bayesian approaches to… evidence and uncertainty" (Clark 2013). PEM conceives of the mind as a system organized in a generally hierarchical structure which constantly aims to predict sensory input and learns from the incongruencies between its guesses and the sensory input it experiences.

To predict the incessant stream of sense-data, the system aims to capture in its own hierarchical structure the statistical structure of reality, or the set of causes which produce the incoming sense-data. Because sense-data can take complex forms and arise from complicated networks of causes (e.g. social interactions and cinema), the mind's hierarchical predictive system must be able to encode high-level knowledge (i.e. store high-level patterns which explain and therefore compress multiple phenomena[5]) to make predictions (Clark 2013). Predictions at a given level of the hierarchical structure are in turn anticipated at

[4] Through movement and motor control, a predictive agent can actively control what its senses will encounter and use that information to confirm, deny, or refine hypotheses about its current situation. This situates the agent in a causal loop from senses to action and vice versa reminiscent of the cybernetic loop described in §1.1. For a particularly good exposition of this dynamic see (Godfrey-Smith 2016)), chapter 4.

[5] By storing the high-level patterns or probability distributions that explain a host of observed phenomena, a predictive system can avoid storing the complete set of observed phenomena themselves. Additionally, this system at certain points need only record the error between sensory inputs and its prediction about those inputs, rather than the inputs in their entirety. From *Surfing Uncertainty*: "As long as there is detectable regularity, prediction (and hence this particular form of data compression) is possible. It is the deviations from what is predicted that then carry the 'news', quantified as the difference (the 'prediction error') between the actual current signal and the predicted one. This affords major savings on bandwidth, an economy that was the driving force behind the development of the [predictive coding] techniques by James Flanagan and others at Bell Labs during the 1950s" (Clark 2016).

higher levels, cycling back. At the lowest level of the predictive structure, the system aims to anticipate sensory data. These predictions provide a probabilistic evaluation of the sense-data sampled at each moment using knowledge represented across all layers.[6]

As described in Clark (2013, 2016) this predictive structure is hierarchical as well as bidirectional. Two-way connections between levels in the hierarchy crucially enable the system to learn from its prediction errors. Downward (or forward, in the statistics literature) connections between layers make successive predictions about, i.e. define a probability distribution over, lower layers.[7] These successive predictions cascade to the bottom of the hierarchy and thus predict the incoming sense data. Information represented at an arbitrary level of the

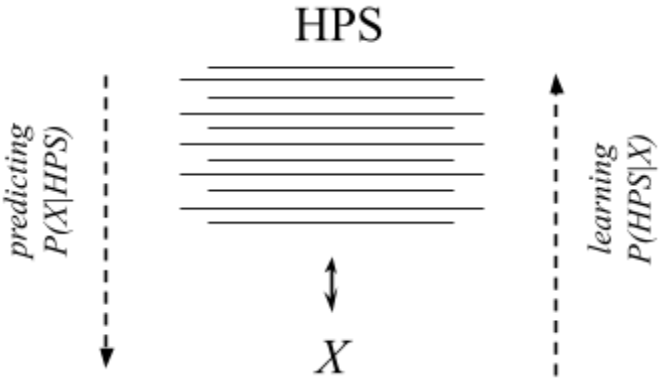

Fig. 1, showing the bidirectional nature of the hierarchical predictive system. HPS refers to hierarchical predictive structure; *X* is sensory data.

hierarchy is passed to and incorporated by successive lower levels such that the lowest level of the hierarchy uses an integrated version of all the system's knowledge to predict sense-data. Upward (or backward) connections allow the structure to propagate information about prediction errors to higher levels of the hierarchy [see *Fig. 1*]. Error information from predictions is used to adjust the hierarchy to better predict future stimuli. This learning process, iterated continuously as the system makes predictions and encounters prediction errors, allows for the evolution of relationships between layers. These relationships, where a layer of the hierarchy acts as a set of prior beliefs for the layers connected to it from below, describe the knowledge—including background information and relationships between concepts retained from previous experience—the system uses to make a particular prediction about sense data. The forward connections between layers direct a flow of concepts used to explain the observed stimulus (as well as concepts at subsequently lower levels of the hierarchy), and backward connections enable updates to the structure and the adjusting of the priors (i.e. concepts) relevant to understanding the stimulus.

One biologically plausible, computationally tractable[8] suggestion for the learning mechanism uses Bayesian principles. Under the Bayesian framework, layers of the hierarchy act as statistical priors for successive layers below it. Prediction involves the formation of a probability distribution over the sample-space of events which might occur in a particular stream of reality (i.e. *umwelt*) experienced by the system. In *Fig. 1*, this distribution is represented generally by $P(X \mid HPS)$, where X refers to sense data inputted to the system and *HPS* refers to the hierarchical predictive structure (and its encoded knowledge); the whole expression can be read as "the probability distribution over X conditioned on the *HPS*." Learning takes the form of Bayesian inference, which requires calculating (or, more likely in a flesh-based system, approximating) the posterior distribution, or $P(HPS \mid X)$. This distribution refers to the distribution over possible hierarchical predictive structures given observed sense-data X, and inference uses this quantity to update the hierarchical predictive structure to best explain X (Clark 2016, 2013).

We are interested in how predictive processing and the Bayesian framework can model a human observer's interaction with a work of art, namely the interpretation of the artwork and the updated understanding of reality the artwork inspires in the observer. The authors are well aware that aesthetic theory is far flung from the areas of cognitive science to which Bayesian models have been applied. The Bayesian approaches from cognitive science which most closely relate to our work here aim to understand modes of cognition apparently distinct from

[6] A structure that is able to predict the sense-data well, i.e. assign a high likelihood to it, is able to "explain away" the sense-data. One must only sample from this probability distribution to generate a "virtual version" of the sensory data (Clark 2013).

[7] Sometimes, the meanings of "forward" and "backward" are reversed in relevant literature, such that "forward" refers to the propagation of prediction error and "backward" refers to the formulation of predictions, as in Rao & Ballard (1999).

[8] For a discussion on the biological plausibility of Bayesian principles see Clark, 2016, p. 39-41.

cognition related to art. We use no experimental regime to investigate our ideas here, and the type of cognition we aim to address appears to be beyond the current explanatory reach of cognitive science research. Rather, we strive to introduce a vocabulary and mode of thinking, synthesized from predictive processing and probabilistic models of cognition, into aesthetic theories concerned with the interpretation of texts, media, and art.

1.3 HBMs as instantiations of hierarchical predictive structures

As described above, hierarchical predictive structures have the basic requirements of prediction (through probabilistic evaluation) and learning (through updates to the structure according to mismatches between predictions and the actual sense-data one encounters in the world). Bayesian principles provide a framework and suite of algorithms to reason about prediction and learning, or inference. What's more, recent research in cognitive science has studied a class of Bayesian models, referred to as hierarchical Bayesian models (HBMs), which learn[9] and reason probabilistically about reality by employing knowledge encoded by the models' structures. Though these models from cognitive science do not assume the fundamental predictive and anticipatory quality of the mind stipulated in PEM, their foundations in Bayesian reasoning allow for their importation to PEM à la *Bayesian hypothesis (Clark 2013)*. If hierarchical predictive structures outline the flows of information to anticipate the environment and learn from incorrect predictions, Bayesian principles provide a framework for contemplating their core mechanisms, and HBMs can be considered a theoretical instantiation of hierarchical predictive structures.

Generally hierarchical in organization, HBMs can define probability distributions over structured symbolic forms such as graphs. Structures of many different forms (e.g. trees, clusters, spaces) can be represented as graphs, each of which has been shown to effectively model certain types of cognition. Such graphical models are able to encode the dependency structure of a system of random variables, including rich networks of causal and correlative relationships. For example, HBMs have been used to represent reasoning for intuitive theories, or "a system of related concepts, together with a set of causal laws, structural constraints, or other explanatory principles, that guide inductive inference in a particular domain" (Kemp and Tenenbaum 2008); (Kemp and Tenenbaum 2008; Kemp, Perfors, and Tenenbaum 2004). The mind is not restricted to reasoning about one particular domain: it consistently utilizes a multitude of intuitive theories, and thus might invoke different networks of concepts (which in turn can be modeled as HBMs) to reason about multiple disparate sense-data streams.

Important to understanding art-encounters through the PEM framework is how HBMs can represent concepts and related principles which contribute to the explanation of multiple phenomena. As an implementation of hierarchical prediction structures, HBMs can effectively represent background knowledge with prior distributions. Critically, high-level priors can be used to represent knowledge, i.e. relationships between concepts, that is relevant in various contexts (Tenenbaum et al. 2011). Such priors can form hypotheses about multiple distinct sequences of sense-data. These hypotheses can be conceived of as cascades of concepts in a rich network of relations that explain observed phenomena.

HBMs are a way of we can express in formal terms the rich processes of prediction and learning. We've seen how they can be used to model intuitive reasoning—one might imagine crafting an HBM which captures the core aspects of engagement and reflection about an artwork. Below we sketch a Bayesian model which encapsulates some of what we consider to be fundamental dynamics in art encounters. This model can be described in further detail as an HBM, which we leave for future work. We stress the abilities that the language and handles of Bayesian reasoning provide us to ponder prediction and learning in the PEM framework. We will elucidate further concepts of HBMs below alongside our discussion of art encounters and PEM.

1.4 Predictive coding and the schema

[9] Learning in complex hierarchical Bayesian models aims to distribute information gained from observing evidence across the structure which aims to predict, or provide a probability measure of, evidence. Learning from evidence requires at the very least calculation of the posterior of the probability distribution defined over the structure and often involves message-passing algorithms. Previous work has shown the structure of concepts governing relationships in a domain can be algorithmically learned (Kemp, Perfors, and Tenenbaum 2004).

We can understand *schema* informally in the way Piaget, Kant, and others have used it: a mental framework which is built and updated through processes of induction and deduction. From psychology and hermeneutics come close synonyms for *schema/ta*, including *framework*, *worldview*, *way of seeing*, *interpretive filter*, and *mental model (Reason 2018f)*. Psychological and perceptual theories of schemas function as a less-formalized cousin to the cognitive-scientific concepts *knowledge structure* (in Bayesian terms) and *hierarchical generative model* (in PEM terms), which is to say the top-down probabilistic models present in predictive and statistical systems.[10]

We'll use "schema" (or "schemas," plural) here as a handle, in the contexts of Bayesian cognitive models and predictive error minimization respectively, to refer to either knowledge structures (encoded by HBMs) or hierarchical generative models (HGMs). This shared terminology is meant to emphasize that though the academic vocabularies of frameworks of PEM and Bayesian cognitive models differ, in this paper's theorized system, HGMs and the Bayesian knowledge structure are analogous; that is, following Hohwy (2013) and Clark (2016), the theoretical conception of PEM used in this paper speculates a Bayesian or Bayesian-approximate predictive structure at all levels.

In other more speculative contexts throughout this paper, "schema" is used as a means of conjecturing potential alignments between the PEM and hermeneutic systems, or as to ways artworks construct and exploit predictive systems, without overstepping into over-speculative mathematical formulations. We hope this stresses that the observations of later sections are not contingent on the specifically Bayesian attributes of cognition hypothesized in §1.3 being the case (nor are the Bayesian attributes contingent on the attributes of art encounters stipulated in §3-5).

1.6 A brief note on art and predictive coding

Predictive coding has been applied previously to visual art, such as in Van de Cruys and Wagemans (2011), Kesner (2014), and Kandel (2016). The focus, however, has always been on visual processing, with higher-level hermeneutic and conceptual processes largely ignored . Kandel cites modern art, for instance, as "dismantling" the literal visual perspective in order to force the brain to "come up with a new logic of bottom-up processing" (Kandel 2016). While such theories are congruous with our own hypothesis, they constitute a mere subset of the types of conceptual disruption this paper theorizes. Here, we depart from these authors by looking at artworks from a high-level, literary-theoretic lens (as opposed to a lower-level, visual art perspective), surveying, first, how hermeneutic processes rely on predictive processing, and second, how works of art and literature actively set up *and then subvert* high-level thematic, narrative, and formal (e.g. stylistic) predictions in their audiences.

§2 Hermeneutics

Textual comprehension requires interpretation, a process of sense-making requiring the extraction of a regularity or logic from textual data. This is true at all levels of textual meaning, from the ground on up: "The conceptual space generated by reading a literary work is a 'cluttered array'" that "consists of colours, edges, forms, and textures that are resolved into *attended Gestalt objects*" (Stockwell 2009). Elements are weighted by relevance and attention is allocated according to the configuration of the text and the schematic interests of the reader.

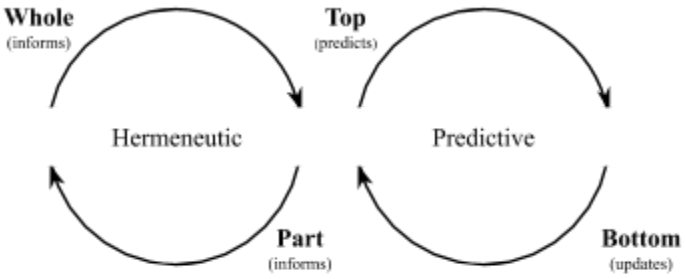

Fig. 2, depicting similarities between hermeneutic and PEM feedback loops.

As argued by de Beaugrande in *Introduction to Text Linguistics*, texts "make sense" in that they contain a "continuity of senses" which are congenial to each other. When texts do not make sense—are nonsensical—they lack such a continuity. The property of "making sense" can be called "coherence," and the configuration which underlies such coherence can be called a "textual world" (De Beaugrande and Dressler 2016).

[10] As an informal example of the ways a schematic worldview can be made up of inferential probability distributions, consider the prototypical conservative worldview, whose perception of high social precarity flows into (and out of) all kinds of risk assessments and outcome probabilities.

2.1 The artwork as puzzle

All works of art, including texts, are puzzles insofar as they must be consciously or unconsciously "puzzled out" in order to yield coherence (in Bourdieusian terms, they must be "decoded" according to a soft grammar, which is to say, "parsed") (Bourdieu 1998). Texts are rarely designed with a "correct answer," though sometimes the authorially intended meaning is presumed to fill this role. Instead, it is the retrieval of any meaningful, non-arbitrary reading whatsoever which requires decoding, be it of the natural language in play (its set of connotations and denotations), the intertextual history of the discipline, the discursively determined values of the tradition within which the work was made and released, etc. In most fiction (and much of nonfiction) the knowledge required to unconsciously puzzle out or "decode" is already possessed by the majority of intended readers and can be applied with little friction. (Indeed, the text is written specifically with such audience-specific legibility in mind.) Such readers are engaged in consistently but unconsciously interpreting both the parts of the "puzzle" and their interrelation, that is, how each part might fit together in a (not "the") coherent whole. New knowledge presented by the work updates the prior conceptions, while prior knowledge influences the understanding of new conceptions. This process involves the recognition of—to name but a few—sequential, logical, thematic, mythological, psychological, and ideological patterns within the work (Reason 2018a). We will consider such encounters, in which interpretation is largely unconscious and self-contained (that is, requiring no reference to outside sources of information), as prototypical hermeneutic encounters, whose structure is reformulated into a predictive framework throughout §2.

2.2 Predictive formulation of a hermeneutic encounter

In §2.1, we described the hermeneutic process as one in which the *observer* (i.e. reader/viewer/audience member) is constantly "engaged in consistently but unconsciously interpreting both the parts of the 'puzzle' and their interrelation, that is, how each part might fit together in a (not 'the') coherent whole" (see *Fig. 2*). Here we will formulate that process approximately in probabilistic terms, representing a hermeneutic encounter graphically through a hierarchy of probability distributions (Sharifian 1997).

Consider a set of events $X = \{x_1, x_2, \ldots, x_n\}$. X is the set of successive events in an art encounter, which may be more narrative-based and determinately sequential (e.g. films and literary texts) or more introspective and indeterminately sequential (e.g. a static piece of visual art). A reader's interpretation of an artwork might refer to the combination of events in set X and a hierarchical knowledge structure C (consisting of concepts and relations between concepts) which are deployed to engage with the set of events X. C should include all relevant knowledge that exists beyond the artwork as well as concepts, relations, and knowledge that apply only in the world of the artwork. We can insert concepts C and events X into a graphical structure, where C and X are nodes in the graph and the edges between them encode their relationships. Using Bayes rule and principles from graphical models, we can use the graphical structure to conceive of a conditional probability distribution $p(X \mid C)$ over the events in X given knowledge structure C *(Koller, Friedman, and Bach 2009)*. A "coherent whole" might refer to a given structure spanning the sets of elements X and C which yields a high value for $p(X \mid C)$, reflecting a high likelihood of the art events given our understanding, or graphical configuration, of related concepts. While engaging with a work over time (e.g. reading sequential chapters of a book, watching a film, interpreting elements of a painting one after another), we constantly make predictions about specific scenes and revise the possible outcomes for future remaining events contained in X. When interpreting a specific event of a work, x_t, only certain parts of the knowledge structure C will be activated: call this subset C', where $C' \subset C$. We also take into account events previously experienced in the art encounter; for example, we can use the events $X' = \{x_1, x_2, \ldots, x_{t-1}\}$ to assess the possible outcomes for the next event X_t, or for the remaining events in the artwork after X_t, written $X \setminus (X_t \cup X')$. Now, we can denote our predictions about event X_t as $p(X_t \mid X', C')$. So as we progress through an artwork, at each timestep we expand, and contract, and phase-shift the activated parts of our knowledge structure C to understand the next event in the piece, the remaining events of the piece, and the piece as a whole.[11]

[11] Each art form and tradition has its own patterns of entropic growth and diminishment—rhyming poetry, for instance, narrows its realm of possibility at the end of each line.

In addition to making top-down predictions *about* art events, we also learn bottom-up *from* art events and update our larger schemas based on information contributed by the art events: we use the posterior distributions over *concepts* given events from the art encounter, e.g. *P(C | X) or P(C'| X'),* to revise the respective distributions over the art events themselves, thereby revising our interpretation of the artwork as a whole[12]—a process markedly similar to that of the "hermeneutic circle" gestured at in *Fig. 2*. As in the hermeneutic circle, this process happens piecewise during an encounter with work, where we calculate posteriors over subsets of concepts from subsets of the data *p(C' | X')*; a scene tells us about a few but not all of concepts relevant to a film. Learning algorithms from Bayesian graphical models, such as message passing, suggest how new observations can propagate throughout and update the structure of concepts. When our knowledge structures of the larger world, or the larger art tradition (and not merely the world of the work at hand), are updated in a significant way, schematic subversion has taken place. We expand expand on schematic subversion throughout §3.

By way of example, we can look at a few of the primary concepts, *C*, in play when watching a Spaghetti Wester: formal genre conventions that intersect with conventions of the era, based on other Spaghetti Westerns, Westerns, and previous films watched by the viewer; historical knowledge about the American West; general knowledge of human behavior put on display; and the physical and operational possibility or "realism"—what we might call the "soft physics" of the real and textual worlds. Crucially, these are merely a handful of relevant informing concepts. Priors will be based also on *all action films watched by the viewer, all Italian films watched by the viewer, all films of any type watched by the viewer, all artworks consumed by the viewer,* etc, where there is decreasing relevance or radial proximity of the concept to the work at hand. Similarly, the viewer's knowledge of human psychology will come into play with respect to the *analogic similarity* of the filmic scenario to learned or experienced knowledge. Relevant concepts grow—but are also constrained, and altered—as the film carries out its runtime.

2.3 Interpretation and intentionality

Insofar as an observer sees a work of art as a form of communication, attentional prioritization will take place on somewhat the same grounds as, or with some degree of symmetry to, what the director intends the viewer to prioritize (since the essence of communication is the ongoing attempt to understand what the interlocutor *means* to convey through an utterance, and not merely what is, technically, said). Audiences search for markers of hierarchical importance in a communication, while artists employ intuitive or else culturally established motifs of emphasis to signal intentionality to observers. Insofar as the artist knows element *m* of the artwork is a marker, the audience knows *m* is a marker, and the director knows the audience knows *m* is a marker (these being the conditions for mutual knowledge), communication is possible. Foregrounding an object, action, or utterance is one default mode in film and visual art to communicate priority, e.g. through volume, prominence, tonal stress differentiation, size, position, or lens focus. Not only are foregrounded events more likely to be naturally noticed, they also direct attention and inform high-level conceptions of communicative priority.

At the sentence level, commas and syntax are two examples of textual markers which enable greater hermeneutic accuracy by signaling intentionality—that is, the specific way the words are intended to be interpreted, or the "continuity of senses" postulated by de Beaugrande. [13] When the meaning of a sentence remains unclear, perhaps due to ambiguous reference or grammatical structure, the interpreter must formulate a guess based on one, dynamics and structure from the real world, and two, assessments of the speaker and his motives. Predictive systems of inference allow us to accurately gauge the intent in the ambiguous sentence *She announced a program to promote safety in trucks and vans* (an example taken from natural language processing). Is the announcement made "in [read: *inside*] trucks and vans"? Did the speaker announce the program *in order to* promote automobile safety, or is it the program which promotes automobile safety?

[12] *P(X|C)*, the probability of *X* given *C*, can be understood as deduction, the process of understanding a particular instance through reference to a general principle. *P(C|X)*, the probability of *C* given *X*, can be understood as induction, the process of understanding a general principle through reference to a particular instance.

[13] Commas and syntax do not "mean" anything; they merely signal and modulate relationships.

In garden-path sentences, such as "The old man the boats," the initial assumptions of a first parse— that "old" is an adjective modifying the noun "man"—must be updated to identify "old" as a plural noun and "man" as in its verb form. In the initial parse, the high probabilistic likelihood of "the old man" signifying an elderly male makes that initial interpretation so strong that the subsequent "the boats" cannot immediately amend it, and a second or third reading is required to verify that the dramatically unlikely minority sense-meaning is, in fact, the one intended.

Finally, at the level of individual word senses, we can theorize a subject who is aware of an approximate probability distribution consisting of the discrete likelihood of each individual sense-meaning—connotative, denotative, rhetorical, and figurative—of a given utterance within its context. Moreover, this probability distribution is known, varyingly and approximately, among the larger population of native speakers. Finally, the distribution, similar to textual markers of priority, is *mutually known*—that is, all members of the group are aware that other members of the group will have similar hypotheses about utterances, this being the step that enables predictable communication, and thus allows us to communicate at all. Language, as we've analyzed it so far in §2.3, can be understood loosely, alluding not just to natural languages but also the languages of cinema, visual symbolism, musical moods or motifs, and so on.

At the phrasal level, predictions are made, for example, as to whether an act of indirect speech (e.g. the veiled threat *It would be a shame if something happened to…*) is meant literally or more suggestively (Pinker, Nowak, and Lee 2008), and both interlocutors in the exchange must share similar-enough models of the phrase's rhetorical probabilistic likelihood within the threatened party's linguistic schemas to communicate. Poetry, as we will see, is engaged, among other things, is undermining and "riffing on" the dominant and thus expected sense-meanings of its language, requiring the poet to have a well-tuned linguistic "metaschema"—a predictive model of readers' predictive models. Similarly, speakers and writers engaged in rhetorical or informative communication must acquire intimate knowledge of how their words and grammars will be parsed by their respective audiences, allowing them to sidestep ambiguities or miscues to get across what they "really" mean to say.

We propose that, as in speech and writing, much of artistic production occurs under similar conditions, whereby an artist's expressions, selections, and framings are guided by the active consideration (i.e. prediction) of how audiences may interpret or anticipate a work from the grammatical to the conceptual levels. However, as we'll see, modernistic and avant works appear to be marked by their specific tendency to undermine or disrupt this anticipatory structure, and accordingly the knowledge structure that informed it.

§3 Art, Prediction, and Subversion

In §2, we considered how a hermeneutic encounter with an artwork could be framed in terms of prediction by a hierarchical probabilistic model. Specifically, we looked at the ways that unconsciously "puzzling out" (i.e. interpreting) a work gives rise to an explanation or hypothesis that gives its parts a coherence of meaning. In much of avant-garde writing, however—as well as in specific literary traditions like modern and postmodern poetry—elements like gappiness, homophonic slip, or grammatical reversal require the reader to actively and consciously puzzle out a work's meaning through repeated readings of the text, or through reference to external sources (e.g. a dictionary, encyclopedia, or fellow literary work). Though it's outside the scope of this paper to define what art "is," we can still observe that there is a strong correlation between how "artsy" (how "sophisticated," "literary," or "avant-garde") a work is perceived as within the cultural field, and the extent to which it actively resists, confounds, subverts, or problematizes the interpretive attempts and procedures of its audiences (Reason 2018b; Noë 2015). That is, the artsiness of art is a property defined by, or at least strongly correlated with, the work's confounding of, and resistance to, *easy* assimilation into observers' schemas.[14] Tellingly, art which does not

[14] Consider that the artistic traditions which gets described in terms of formal advancement or dialectic—that is, the pushing of material, tonal, formal and categorical boundaries in experimental and avant-garde fields (Gordon and Poggioli 1968; Bensman and Gerver 1958; Crane 2015; Clignet and Crane 1988)—can also be described in terms of the continuous subverting expectations of observer assumptions, roughly equivalent to subverting the work's operational assumptions. Here, the avant-garde's trajectory (or "progress") can be described not as the primary goal of avant producer but as a byproduct of a continuing quest to disrupt the

complicate or subvert audience expectations is perceived as formulaic, predictable, boring, or trite, and is accorded less prestige as a result (Martorella and Crane 1988, 12); (Bourdieu and Johnson 1993).

3.1 Schematic subversion

> That is the danger with Kafka. Just when you think you know him he makes a sharp turn and you end up facing a wall.
>
> —James Nulick, *Valencia*

Stockwell in *Cognitive Poetics: An Introduction* summarizes "schema poetics," a subdiscipline of cognitive poetics concerned with readerly context and interpretive lenses (within our framework, a reader's probabilistic assessments of the world). Stockwell divides discursive modes into those which are *schema-preserving*, *-adding*, *-reinforcing*, and *-disrupting*. Everyday discourse is typically schema-preserving or schema-adding, in that it tends to conserve and apply its interlocutors' worldview. Literary and artistic discourse, meanwhile, tends toward the disruptive (Stockwell 2005).

In analyzing the information content of discourses (from the everyday to the literary), Stockwell cites Robert-Alaine de Beaugrande, who in 1980's *Text, Discourse, and Process* carves categories of discursive information into first, second, and third orders. *First-order informativity* entails low novelty and low surprise, and is therefore schema-preserving. *Second-order informativity* presents more "unusual" data and helps "develop schematic knowledge by accretion," adding nuance or detail but stopping short of significantly altering high-level priors. *Third-order informativity*, finally, involves "highly unlikely" or unusual data with respect to the reader's incumbent schema, and can represent such a disruption as to lead to its thorough restructuring.

For example, we might have a predictive structure partway through a novel for "what kind of person" a literary character is. This schema can be represented by a hierarchical model, resembling the one presented in §2, which encodes in its structure the knowledge useful to understanding the character and predicting their future actions and outcomes. Said schema will require non-trivial revision should we come upon surprising information about that character's actions or motivations later on in our experience of the text (de Beaugrande 1982). Actions disjunctive with our gestalt understanding, or new information which alters our understanding of a character's motivations, may reshape, retroactively, our understanding of the character's previous actions.

We'll refer to *forced hermeneutic revision* as a type of schematic subversion (schema-disrupting, to extend Stockwell's terminology)—the process by which texts and artworks methodically set up audience assumptions through coherence, as outlined in §2 and §2.2, only to systematically undermine many of those assumptions later in the work. A variant of this strategy we term *opportunistic hermeneutic revision*, in that it capitalizes on implicit audience assumptions about how a category of works "work" (e.g. what is expected of genre fiction, or of music, or of poetry) that have been culturally built up and are thus capable of being subverted forthright. A common effect of poetry, for instance, is the subversion of dominant sense-meanings of a word or phrase in favor of more minority interpretations (see §2.3 Interpretation and intentionality). From (Pierre Bourdieu, Bourdieu 1985): "'Pure' poetry appears as the conscious and methodical application of a system of explicit principles which were at work, though only in a diffuse manner, in earlier writings. Its most specific effects, for example, derive from games of suspense and surprise, from the consecrated betrayal of expectations, and from the gratifying frustration provoked by archaism, preciosity, lexicological or syntactic dissonances, the destruction of stereotyped sounds or meaning sequences, ready-made formulae, *idées reçues*, and commonplaces."

In general, schematic subversion involves the presentation by the work of third-order informativity, that is, information which has high novelty and surprise given audience schemas.[15] Expectations which are forcibly or opportunistically subverted include understandings based on genre, tradition, author, cover, synopses, reviews, social information, etc., and stem from a combination of previous

expectations of their audience. There is some precedent for this: Dada has been described by Duchamp, for instance, as the "denial" of established interpretive approaches (Duchamp 1956).

[15] Schematic subversion can also be understood in the Murray Davis sense of interestingness, where interesting information is that which updates the assumption-ground (set of assumptions) of its audience (Davis 1971).

experiences and learned information, both within the space of the work and outside it.

In other words, effective artists are keenly aware of our predictive structures, and frequently leverage such understandings to not just maintain attention or communicate with clarity and economy, but also to create suspense and upend conceptions. One mode, associated with "high" or "avant" art, tends toward subverting our generic, formal, and textual expectations of a work; in certain subfields of visual arts, ideological subversion is a primary aim (Reason 2018d). Within "low" or popular art, these subversions are more often narrative-based than structural, taking the form of plot twists and comedic relief. Such works cause us to fall into predictive traps, into understandings or anticipations of the artwork or world of the artwork which are shown to be misguided. Within the framework of cognitive poetics, we can say that art is schema-disrupting, and much of its value to audience members derives from its third-order informativity.

3.2 The artwork as a joke

The conception of artworks as schema-disruptive presented above bears commonalities with the theory of humor presented by Hurley, Adams, and Dennett in *Inside Jokes: Using Humor to Reverse-engineer the Mind* (2011). The essence of a joke, in the view of Hurley et al., is that its teller surreptitiously introduces a certain epistemic commitment, then reveals it to have been mistaken. When we experience humor, we are led down a "garden path" of a covertly introduced, mistaken assumption which is then revealed to us via the punchline (Hurley, Dennett, and Adams 2011, xi–6).

Hurley et al. use the computer science metaphor of debugging, and make an evolutionary argument as to its origins. "Mother Nature—natural selection—has [stumbled upon humor as an incentive for] our brains to do all the tedious debugging that they must do if they are to live dangerously with the unruly pile of discoveries and mistakes that we generate in our incessant heuristic search" (Hurley, Dennett, and Adams 2011, xi).

The philosopher Henri Bergson presents a similar analysis of the humor inherent in tripping and other physical or social awkwardnesses. Bergson argues:

> …when people are too trapped in the automaticity of their mechanical movements and when these are insufficient in dealing with the environment at hand, a comical situation presents itself. Bound by the habits of movement, people sometimes forget to adjust for new terrain or unexpected obstacles, or they get so accustomed to their standard environment, they expect the body to do all the work intuitively. (Wampole 2015)

When one's predictions fail in an art encounter, it is much like Bergson's jolt: out of the automaticity of interpretation (as well as the schema responsible for the misinterpretation in the first place) and into a state of cognitive arousal [*see* §4.1, Art as superstimulus, & §4.3 Surprisal as situationally valenced].

3.3 Charged subjects and formal gestures

It is our belief that many artists, rather than making a work which is "realistic" in itself, choose "representative" scenes, characters, and images to be the building blocks of their works. These elements are somehow archetypal or demonstrative of an underlying reality, sometimes physical but more often social, psychological, or cultural. In this, we have the support of reader-response theorist Wolfgang Iser: "No literary text relates to contingent reality, but to models or concepts of reality, in which contingencies and complexities are reduced to a meaningful structure" (Duckworth 1979).

Readers and viewers therefore encounter "charged" scenarios—data points which, because they represent, to the artist, rich, complex phenomena in the real world, tend to be highly meaningful, intended to be contemplated at some length after the experiential fact (and which often naturally are, given their difficulty, complexity, and interestingness). From a humanistic perspective, the subversion of these scenarios is extra meaningful; the subversion is not "arbitrary" but tethered to a world and acts on the array of related schemas which might generalize to other works and domains. In §4.4, we'll further consider how these "charged" or representative data points might be perceived as higher-importance, highly weighted inputs to a mind's probabilistic model of reality.

In humanistic, cartographic art, schematic subversion is typically done of a charged subject, one which relates to the real world and has a significant bearing on its audience's schemas, e.g. their understandings of social psychology. In more formally oriented works, the subverted subject might relate to specifically linguistic or spatial expectations, such as the probabilistic sense-meanings of a word (§2.3) or else the unspoken rules and boundaries of a discipline. Artists like Duchamp and Cage epitomize the latter approach, while pure poetry is chock-full with the former. Both constitute foundational priorities of modernist and twentieth century art practices more generally.[16]

From James Schuyler's "The Morning of the Poem":

(1) *Force, fate, will, and you being you: a*
painter, you drink
Your Ovaltine and climb to the city roof, "to
find a view," and

(2) *Wings in fierce blue delphinium depths I think*
About those two blue jays like me, too
chubby, and Baudelaire's skull

In the first excerpt, "drink" leads to predictions of alcohol that are revised (or doubled) by "Ovaltine." In the second, "too" is initially likely to be understood as meaning "as well," an assumption which is revised (or doubled) by "chubby" to meaning "overly."

From Bernadette Mayer's "Synesthetes at the Writers House":

I'm pleased to announce
that staying at the Writers House
is like living under a multi-colored apple tree
in winter;

Here, the divergence of figurative connotations between a Writers House "like living under a multi-colored apple tree" (with its suggestions of bounty, fecundity, beauty, and shade) is subverted in the next line by the suggestions of that same tree in winter: barren, providing no shade, ugly in its nakedness. Mayer's writing, like so many poets and artists, has been praised for its subversion and manipulation of reader interpretation: "[She] postpones interpretation, perhaps forever, in an attempt to chip away at both her and her reader's compulsion to know where the writing is going [next]," writes Maggie Nelson in *Women, the New York School, and Other True Abstractions* (Nelson 2007)**.**

Lastly**,** it is worth drawing attention to the concept of the "turn" or *volta*, in theories of poetry. This is the moment in the poem in which its tone or rhetoric pivots noticeably into another register; it is often seen as the heart of the turning work, the revelation which reveals its center.[17] In other words, what comes after, or with, the volta retroactively modifies our understanding of what has come before, and often goes so far as to directly subvert the understanding which the poem has been building up theretofore.

3.4 Hermeneutic revision and the effect idea

Schematic subversion is an effective way to not just adjust but also bare or expose expectations (e.g. bias, preconception, worldview). *Effect ideas* are artistic mechanisms of action—by a work, onto an audience member—in which the baring of the assumptive ideology (i.e. schema) conveys with it valuable information, or else poses valuable questions, about the world. A "form of philosophy," effect ideas exert themselves through the reader *watching himself watch the text*, predicated on cognitive-predictive self-awareness throughout the art encounter. Through this self-watching, the reader comes to understand more about not just art, reality, and the world but about the schematic self—the set of probabilistic expectations brought to the encounter—through said schema's enactment on the work (Reason 2017a). Frequently, though not always, the effect idea hinges on a reader's confused, ambiguated, or otherwise subverted reading of a charged subject.

In this way, the artist catches us in our prejudices and assumptions, and suddenly bares them—which is to say he *flashes* us with our usually invisible ideologies, suddenly

[16] Again following Bourdieu, we can speculate that this is due in part to the difference in audiences between avant and non-avant work. In the 20th century, the visual art field especially but also more largely the fields of theater and "high" literature became increasingly "autonomous": self-sufficient and "insider," with a large percentage of its consumers themselves being invested players (artists, curators, collectors) in the field, as opposed to a more disinterested observer. As fields gain autonomy from the larger society and market, they increasingly devote attention to the specific methods, techniques, and approaches of the field (since these topics will be of greater interest to "insiders" as opposed to "outsiders.")

[17] "We could say that for the sonnet, the volta is the seat of its soul." (Levin 2001)

and with recognition. Like Bergson's "jolt" describing a tripping pedestrian (§3.2), the automaticity of interpretation has been disrupted, forcing an evaluation of what went wrong. The attentive audience member gains an awareness of his own, revealed schema, thus we can creatively classify the effect idea, following Stockwell's system of naming, as *schema-baring*.[18]

Though the effect idea plays an important role in all mediums, the visual fine arts and literary fiction especially encourage this category of response. Their consumption involves prolonged pondering and self-evaluation when faced with the art object. Moreover, literary and gallery audiences have been trained to treat these mediums this way (that is, self-reflectively), and in response, the mediums' works are created with such a treatment in mind. *(Grietzer 2014)*): "Art is more 'artsy' the more [the hermeneutic] process bolsters its intended impact." Tellingly, conceptual art hinges on the effect idea, using the subjected experience as a way to communicate—allowing for an observer experience that is less pre-determined or 'railroaded' compared to alternative models.

Still, effect ideas crop up in unusual places, including the oeuvres of artists perceived as "middlebrow" or lacking in cultural prestige. Spielberg employs a predictive lure in the openings of many of his films, in which the audience is cued into misapprehending the opening shot. In *Close Encounters of the Third Kind*, Spielberg leverages the audience's trailer-cued expectations (that the film's subject will be extraterrestrials) by showing two bright beam-lights on a dark screen. What appears at first glance to be a spacecraft is revealed seconds later to be a jeep in a sandstorm. In *Jurassic Park*, Spielberg uses audio production, rustling leaves, and ambiguous shot framing in order to convince its audiences that they are looking at a dinosaur—only to reveal that the object in question is a large transport truck. In both cases, what is initially presumed to be a mysterious or monstrous "other" is revealed in fact to be an instrument of man, an effect which carries with it an implicit set of ideas.

[18] For clarity, Stockwell's original carving consisted of schema-preserving, schema-accreting, and schema-disrupting modes of discourse. To this trio we have additionally proposed the "schema-subverting" (§3.1) and "schema-baring" modes (§3.4).

§4 Characteristics of schema-subversive art

In §3, we theorized the *schema-subversive* (both forced and opportunistic) and *schema-baring* functions of art objects within the predictive framework introduced in §1 and 2. Here, in §4 and 5, we speculate on other possible dynamics between art and an audience's schemas.

4.1 Art as superstimulus

Following Hurley et al. (2011) on humor , we propose that art is, among other functions, a kind of higher-level cognitive superstimulus culturally evolved to target humans' innate predictive structure. Agent arousal correlates with the properties of high perceived relevance or precision as a Bayesian input, derived, respectively, from the work's (perceived) topicality and the author's (perceived) credibility. Like a joke, which is tailored to guide listeners to a specific interpretation of events only to pull the rug out (Hurley, Dennett, and Adams 2011), art is tailored to target existing compressions in a subject's schemas. Where classical art often reifies or activates familiar patterns, e.g. patterns used for object recognition (Evans 2019), contemporary works often exploit and subvert regularity observable in the real word. In either case, the artwork provides an intense encounter between a subject and schema, on one side, and the art object with its highly compression-prone or compression-breaking information on the other.[19]

4.2 Truthiness and adherence to model

In an art encounter, our mind updates its inferential models about the world with respect to the work's perceived accuracy, an assessment made by the model itself. The observer's schema acts as a "check" or arbitrator on its own incorporation of the artworks' worldview (dynamics and concepts);[20] when the cartography of the work is too implausible in the eyes of an apprehending schema, it may be dismissed entirely. This is to say that concepts learned directly from personal experience, and indirectly from

[19] The aesthetic theories of John Dewey, especially 1934's *Art as Experience*, is a guiding force here, stipulating as it does that art exists not as an inherent property of a material, but as a dynamic exchange or interaction between an activating viewer and activated object.

[20] More accurate within our PEM framework, we can say that the work's accordance with the viewer's schema plays a key part in guiding the inferences of, or updates to, the HGM, based on the art encounter.

outside sources, are used to assess the likelihood of a work's worldview as conveyed through its components.[21] The artwork's "truthiness" as estimated by the viewer can be understood as its precision, or reliability (Clark 2016). In this way the schema can be understood as a gatekeeper to its own revision: only stimulus surpassing some level of intelligibility and precision for a viewer will be able to interact with the viewer's schema and incite revision. One consequence is that information which fits closely with an existing schema but poorly with a ground-truth reality is perceived by that schema as more, rather than less, likely to be the case.

Art that presents worldviews or models of reality that are congenial with the observer's can be termed *resonant*; art that is presented by a source who we deem authoritative is termed *credible*.

4.2.1 Resonance

Resonance is a term used primarily by non-scholarly readers (Stockwell 2009), and the vagueness of the term as used in non-scholarly discourse fields has led to its being avoided in academic discourse.

As we use it here, to *resonate* is to oscillate in such a way that (figuratively) "a sympathetic oscillation occurs in a similar nearby structure," (Reason 2018e) and for most readers, engaged in a highly personal encounter with a text, the original oscillating structure is the text, and the sympathetic body, in whom resonant vibrations occur, is the reader's body of personal experience or worldview—their schema of the world.

While resonance is one of our key checks in art-based learning, it is highly susceptible to confirmation bias. Radical updates moving toward truth may be dismissed for their implausibility according to the observer's standing world model, while minor updates moving toward objective falsehood might be deemed more plausible within the standing schema. Moreover, the kinds of transformative texts that we look to in literature often work off mental models of the world so distant from our own that they ring alarm bells, but it is precisely these works which are so valuable to learning (Reason 2018e). This situation is compatible with the understanding of psychosis through "pathological priors," beliefs which while damaging or patently untrue, nonetheless are sought out by a dissonance-minimizing system to reduce entropic uncertainty (Carhart-Harris and Friston 2019).

4.2.2 Credibility

We appear to treat artistic, literary, and cinematic works simultaneously as communications, where an interlocutor's intent constitutes valuable information, and as maps or models of reality (the "cartographic function"), which is perhaps why concepts like authenticity and sincerity are among the most contested and consecrated in literary-artistic production. Insofar as we see artworks as reality-updating inputs to the brain, our trust in and opinion of the source of information will be highly valued. In rhetoric, "credibility" is important as a marker of the accuracy of presented information, or the efficacy of an action proposal.

One significant innovation in fiction was the pointed use of an unreliable narrator. Narrators are considered "unreliable" when their reports significantly diverge from what we, the readers, are led to perceive as the "reality" of the textual world. We can understand this, similar to how pseudonyms disrupt readers' hyperpriors of *author identity*, as a means of further destabilizing the interpretive process (see §5.3 Hyperpriors).

4.3 Surprisal as situationally valenced

The disruption of predictive response, i.e. surprisal, has been theorized to cause arousal in the subject due to its signaling "important changes in the environment" which require acting on (Mandler 2003). This arousal may be neutral in that the situational context of its occurrence affects whether it manifests for the subject as a positive emotion (e.g. interest) or a negative one (e.g. stress); in other words, the arousal is *situationally valenced* ((Van de Cruys and Wagemans 2011).

In an art encounter, the situational context contains both the generic set of expectations specific to the artwork as well as an implicit understanding that no action need be taken in response to the environment of the work. On some level, even as much as one "mistakenly" views an artwork (especially a fictional narrative or world) as real-world

[21] Following memory research, we can think of concepts derived from personal experience as *episodic*, and concepts learned second-hand as *semantic*. (Tulving, Donaldson 1972).

data to populate one's models, the artwork is still cordoned off as a sense-datum of low-risk or minimal consequence (and which, opportunely, might reflect properties or phenomena of reality). We propose that in an art encounter, the observer is immersed in events and predictive work within a controlled, low-risk "sandbox," causing predictive updating (i.e. inference work) to be low-anxiety, even pleasurable, rather than stressful (Reason 2018d).[22] This low-risk quality is reinforced by the fact that predictive events and failures in art experiences are typically either asocial (updates happening in one's private thoughts) or shared and automatic (the collective gasp of an audience at a horror twist)—thus they further lack the social cost of having one's poor calibration publicly revealed.

An alternative theory of art arousal, presented by Van de Cruys and Wagemans (2011) and taken up by Kesner (2014), argues that the prediction-defying nature of a work creates an initial displeasurable dissonance that can be resolved—pleasurably—when the work's deeper structure (or grammar, or logic) is eventually discovered (Van de Cruys and Wagemans 2011); (Kesner 2014). Phrasing this within our PEM framework, we might say that an observer enters an art encounter with a set of activated interpretive concepts, *C*, judged appropriate to the artwork. When those concepts (and therefore the larger schema) are defied (as they inevitably are, in non-superfluous, meaningful works), the observer must propose different hypotheses by which to understand the work, and the eventual discovery of an appropriate schema produces subjective pleasure.

This alternative hypothesis finds support in Schmidhuber and Friston, who both ascribe pleasurability to discovering efficient, instrumentally accurate explanations (or "compressions") of previously unpredictable stimuli (K. Friston, Kilner, and Harrison 2006; Schmidhuber 2010; Schwarz 2013).

4.4 Art as "weighted" or high gain

Art encounters, while inevitably "noisy" at the cognitive level (*one is distracted, the mind drifts…*), are a culturally special practice characterized by low-distraction environments (gallery white walls, dimmed theaters, a spotlight) in which the subject's attention is focused primarily on the stimulus at hand. Attention, as Clark advances, increases the *gain* (i.e. weighting) on prediction errors, increasing in turn the learning potential of an encounter, that is, the significance of its effect on a subject's schema (Clark 2016). After the fact, art encounters are frequently contemplated at length, often socially or textually (e.g. in critical reception) in discursive scenarios not dissimilar to Hohwy's model of *introspection as inference* found in 2013's *The Predictive Mind* (Hohwy 2013).

Further, we can see art as a culturally special practice in which not only are large amounts of attention focused, and focused specifically on the inferential work of *figuring out* and *sense-making*, but there is additionally a credibility (in the sense of §4.2.2) to both the specific artist and the fact of an encounter itself (artworks being culturally special) which might assign high precision to error signals propagated during the art encounter.

§5 Further applications to aesthetic theory

The predictive model of cognition also casts light on a number of related terms from literary theory and aesthetics. Here we'll frame some of these concepts in a predictive framework in order to illustrate PEM's explanatory power within established humanities frameworks.

5.1 Accessibility

When we say a work (or category of works, e.g. conceptual art) is inaccessible to the general public, what we mean is that the ratio of familiarity to foreignness, of predictable to unpredictable, is so low as to make the work unassailable by the observer. There is no puzzling to do because there is a dearth of priors from which to puzzle, and the reader is left, in the words of (Csikszentmihalyi and Robinson 1990), "on the outside, unable to interact with the work." In an environment dominated by noise, the observer is left without a starting place, without any priors or constraints on whose basis he can expect certain characteristics over others, or form any predictions at all

[22] Speculatively speaking, it would be *surprising*, given apparent bodily physiological changes in response to low-stakes vs. high-stakes situations (anxious vs. relaxed states), if there were not a corresponding difference in the feeling of arousal which results from prediction errors while in each of these states. We can imagine that predictions about the relative danger of a prediction error would even be baked into the prediction structure itself.

about the work. This phenomenon might reflect an inability of the viewer's hierarchical predictive models to provide meaningful hypotheses about the artwork.

5.2 Ideal vs. implied readers

The "ideal reader" as a literary-theoretic concept is often used (implicitly) to mean a reader with an identical significative and interpretive schema to the author, such that communication between them is, so the thinking goes, losslessly transmissited. But this ideal reader (as Wolfgang Iser astutely points out) would find the author's text entirely "superfluous," having nothing to learn from it. Instead, Iser points to an implied reader, who "embodies all those dispositions necessary for a literary work to exercise its effect" (Iser 1978). In a schematic frame, we can understand this implied reader—ideal in his own way—as possessing beforehand the very priors and assumptions which the author assumes in crafting a situated sequences of meanings and subversions in his writing the text (assumes both conceptually and also with respect to the probabilistic "sense-meanings" in the language, *see* §2.3 & glossary).

5.3 Hyperpriors

Hyperpriors in PEM systems are high-level priors employed in an interpretive situation, such as an art encounter. Pervading hyperpriors in our cognition can include not just time and space—e.g., following (Clark 2013), "that there [can] only [be] one object… in one place, at a given scale, at a given moment"—but identity —following (Carhart-Harris and Friston 2019)), our "narrative of self." Hyperpriors, in our probabilistic formulation of art encounters, are high-level priors that constrain the space of possible subnetworks of concepts and their relations (Clark 2013; Swanson 2016). As such, hyperpriors invariably influence the selection of a dominant hypothesis for given sense-data, and all hermeneutic work is guided and biased by them. Hyperpriors tend towards being "fixed" as opposed to "fungible"—for instance, hyperpriors are less susceptible to revision, both generally and in an art encounter, than hypotheses regarding the work's meaning or content. However, many artists do, in fact, attempt to misguide audiences with faked or absent hyperpriors, such as by masking their identities through pseudonyms, or presenting their works in non-art contexts.

5.4 Defamiliarization

A well-known phenomenon in psychology is the cessation of full awareness of familiar stimuli. The phenomenon has been called "compiling" by Herbert A. Simon, "tacit dimensionality" by philosopher Michael Polanyi, the "ready-to-hand" by Martin Heidegger, and "driving on autopilot" in casual parlance (Ekman 2013). In the terms of the Russian Formalist school, it is the difference between *recognition* and *seeing*, where to recognize is to perceive in a minimal, peripheral way. Under the PEM framework, we can understand this as schemas "explaining away" well-integrated stimuli. Where high-novelty information, or significant clashes between our expectations and sense data, earns more awareness in this model, low novelty or schema-congruent sensory data is allocated less space in the consciousness field.

To the Formalists, a crucial aim of art was defamiliarization, where the everyday and banal (or "ready-to-hand," or "pre-compiled") is presented in a way which distorts it into newness. Audiences appreciate a fresh sight of what was previously merely recognized. Often, the defamiliarized subject is not immediately recognizable for what it is; only when the mind connects the defamiliarly presented with the familiarly known, an analogic link is created between the two which upcycles into new models, or interpretations, of the familiar.[23]

An adjacent concept to defamiliarization, taken from psychology, is *cognitive disfluency*, which describes effortful attention to a stimulus (in contrast to automatic or effortless attention). Disfluent experiences, such as the art encounter, have been found to "improve syllogistic reasoning and reduce reliance on heuristics" (Hurley, Dennett, and Adams 2011). Tellingly, both awkward situations (*see* §3.2) and avant artistic encounters are characterised by the disruption of automaticity in favor of more disfluent hermeneutic or affective states.

For discussion of the design of intentionally fluent experiences, see §5.8, "Design as the antithesis of art."

[23] Jonathan Richman, *ArtNews*: "[Dalí's] paintings helped me find my way with their tone of foreboding and mockery but also a sense of wonder at the universe itself and what Aldous Huxley once wrote about something else: 'the sinister otherness of familiar things.'"

5.5 Ambiguity

Ambiguous states are those which lack a dominant explanatory hypothesis. Often, multiple alternative meanings are visible in what is known as polysemy; other times, the allusion of the referent is unclear and no hypotheses generated by an observer's schemas earn a high enough probability to be considered likely.

5.5.1 Temporal decay in probability mapping

For further context, see §2.3, Interpretation of intentionality. Shared probability distributions of meaning and reference allows artists and audiences to communicate. Over time and across cultures, this symmetry between interlocutors decays, leading to an ambiguity of reference wherein no dominant meaning—or a different meaning than before—presents itself. As a result, the work becomes less accessible. The hook of Nina Simone's "Mississippi Goddam"—"And everybody knows about Mississippi, goddam!"—once unambiguously ("everybody knows") referenced the state's racial regressiveness in the midcentury. The structural racism sense of the allusion "everybody knows about Mississippi" might still retain a plurality of likelihood if used today, but this sense is no longer clear-cut.

5.5.2 Ambiguity of priority and grounding

When an artist is ambiguous or counter-intuitive in the presentation of an object or occurrence such that it is unclear where attention ought to be paid—e.g. by placing signals in the background, or among noise—our predictive models of perceptual triage are questioned. If the artwork's ambiguities accurately map real-world ambiguities, then we are updating our predictive models with valuable nuance (that there is not a single, dominant interpretation but many). While we may not be able to make stronger predictions of reality, our confidence in probability distributions (a prior on the distributions) has been altered, which will change the structure of future predictions (Reason 2017b). The effects of ambiguous artworks, in this sense, are not dissimilar to the effects of psychedelics as outlined by Carthart-Harris & Friston in "REBUS & the Anarchic Brain," where REBUS refers to the relaxation of priors under the influence of psychedelic compounds (Carhart-Harris and Friston 2019).

5.6 Predictive coding and music

Music is the arts discipline which has been most interfaced, to-date, with predictive models of cognition. Various studies have linked the anticipation of melodic patterns to lower-level auditory coding. Researchers have noted that unfamiliar music is less dopaminergically rewarding to listeners, whereas more familiar music, in which listeners can actively anticipate coming notes and chord shifts, is more rewarding (Salimpoor et al. 2011, 2015; Dura 2007); similarly, chance compositions have been found to be less rewarding or pleasurable than those which work within more predictable structures. These observations can be applied in thinking about why most popular music operates off a small bank of chord progressions, why records get "worn in," and why atonal and achromatic music is less appealing to casual listeners (Reason 2016). It may further help explain rhyme schemes, chorus/verse structure, and the concept of "difficult" versus "easy" listening.[24]

In musicology, a relative of predictive processing models has emerged in Eugene Narmour's implication-realization model of melodic expectation (Gjerdingen and Narmour 1992; Narmour 1990, 1992).

5.7 Mass culture and mutual error minimization

Clark (2013) frames the phenomenon of media-enabled communication (e.g. references to fictional situations and characters, or description through analogy to mutually known media) in terms of collective prediction error minimization:

> Using a variety of tricks, tools, notations, practices, and media, we structure our physical and social worlds so as to make them friendlier for brains like ours. We color-code consumer products, we drive on the right (or left), paint white lines on roads, and post prices in supermarkets. At multiple time-scales, and using a wide variety of means (including words, equations, graphs, other agents, pictures, and all

[24] "Difficult" records are understood to be those which take time and patience to "wear in"; the internal structure or organizing logic proves difficult to grok, or else so far from familiar structures that they cannot be laterally understood. "Easy" records, a term often used pejoratively in referring to pop music (see "easy listening"), are experienced as having an "instant" or first impression charm.

> the tools of modern consumer electronics) we thus stack the dice so that we can more easily minimize costly prediction errors in an endlessly empowering cascade of contexts from shopping and socializing, to astronomy, philosophy, and logic (Clark 2013).

This mechanism as it applies to mass-consumed art sees support in media studies thought (e.g. vis-a-vis globalization), where ink has been spilled over the ways a mass medium like film creates more homogeneous (and thus, implicitly, predictable) cultural rituals, for example in romantic norms. Thus the literary-theoretic concept of texts as "models" (Iser 1993) takes on further meaning (as not merely descriptive but also normative).

5.8 Design as the antithesis of art

Philosophers of art such as Alva Nöe (2015) have previously defined art by its distinction from design, similar to the ever-present contrasting of craft and art in 19th and 20th-century discourse. Insofar as art's subversion, confounding, resisting, and problematizing of interpretation can be understood as purposeful illegibility and surprise, the world of design is characterized by its adherence to the principle of supreme, low-surprisal legibility. Successful design appears to require insight into audience (or "user") schemas and procedures equal to that of successful art. However, this insight is used to cue and prime users of designed technologies and interfaces to ensure smooth experience and intuitive navigability. Much of our designed world has, by design, become thoroughly familiarized, backgrounded into our lives and processed unconsciously (i.e. unproblematically; see §5.4 on "explaining away" sense data) (Reason 2018c; Noë 2015; Reason 2018b).

As Nöe writes:

> A designer of doorknobs makes a simple artifact, but does so with an eye to its mesh with this larger cognitive and anthropological framework. When you walk up to a door, you don't stop to inspect the doorknob; you just turn it and go right through. Doorknobs don't puzzle us. They do not puzzle us just to the degree that we are able to take everything that they presuppose—the whole background practice—for granted (Noë 2015).

5.8.1 Temporal decay in considerations of "art"

Moreover, as any one artwork fades into the horizon line of history, it can cease to register as art because the schema-subverting and schema-confounding effects which once characterized it are temporally and culturally (i.e. chronotypically, following (Bakhtin 2014)) hyper-specific and thus no longer in play for contemporary audiences. Works which we grow up with can rarely surprise us, are difficult to anew. Nöe again:

> Very often we find ourselves admiring old masters, for example, more or less solely for their decorative aspects, or because of their supposed historical significance or monetary value, or perhaps because they exhibit virtuosity in craftsmanship. And so of course it seems implausible that we admire works of this sort because of the way they subvert or undercut or abrogate the authority of what is normally taken for granted. After all, that's just not what these works do for us, at least most of the time. They have expired. Or stopped being artworks (Noë 2015).

5.9 Future directions?

This paper cannot provide a full interfacing of literary theoretic, art historical, and aesthetic frameworks with theories of predictive error minimization. We'll close only by laying out a handful of concepts which jump readily to mind as potentially congenial with such theories : *critical reagents*, *free indirect discourse*, *gappiness*, *genre*, *grammar*, *indirect speech act*s, *interestingness (as defined by Murray Davis)*, *irony*, *the logical consistency of fantasy worlds*, *narrator reliability*, *punning*, *red herrings*, *representation*, *rhetoric*, *rhyme*, *suspense*, *unreliable narrators*, *values hierarchies*, and *visual patterns*.

GLOSSARY

art encounter: following John Dewey's *Art as Experience*, a dynamic exchange between a viewer and an art object.

Bayesian inference: a statistical inference method in which Bayes' theorem is used to update the probability of a hypothesis of a model using prior beliefs and new evidence as it becomes available.

bidirectional: referring to the way, in multilevel predictive systems, bottom-up sense data and top-down *schemas* intertact.

cartographic: referring to the way an artwork "maps" or models reality, e.g. as prominently found in the social realism of 19th century fiction.

compressability: regularity of a data set, either inherent or with respect to some pattern identifiable or predictable by a subject's *knowledge structure.*

hermeneutics: the practice of interpreting texts, used loosely in this paper to mean the interpretative work necessary in any art encounter.

hierarchical Bayesian model: a class of Bayesian models which learn and reason probabilistically about data by employing knowledge encoded by its structure, i.e. probability distributions connected in a hierarchical graph. HBMs are an instance of predictive schemas.

hierarchical generative model/s (HGM): Clark's carving of the predictive *schema,* which may or may not utilize Bayesian learning.

hyperpriors: a systemic prior, tending toward the abstract and fundamental, in subjects' predictive models (see Hohwy 2013, Carhart-Harris and Friston 2019, Clark 2013, Swanson 2016).

interestingness: a subject-dependent property of information occuring when said information allows an observer to predict or compress the world either more efficiently, or more accurately (see Schmidhuber 2009, Davis 1971). As Davis argues, schematically redundant or random information is typically understood as uninteresting, while interesting theories subvert, compress ("help cohere"), or complicate and help revise one's models.

knowledge structure: the set of concepts or priors and their relationships constituting an HGM, which guide its predictions and probabilistic evaluations of the environment.

noise: incoming information left out of models as too high entropy (low *compressability*) or irrelevant (and thus excluded from conscious attention).

precision: the reliability or certainty of information in an inferential system, as estimated by that system.

prediction error: a measure of distance between the predicted value (by a model) and the actual value. In practice, the distance between a predictive agent's expectation of certain evidence according to their priors, and the actual perceived evidence.

predictive coding: a framework in sensory domains like visual processing that the brain actively anticipates, rather than passively receives, incoming sensory information

PEM: prediction error minimization, a cognitive science framework which relates prediction in agents (see *predictive coding)* to the ideas of *hierarchical generative models* and *probabilistic inference*

observer: a generic term referring to both the "viewer" (of film or visual art theory) and "reader" (of literary theory) allowing for a liberal discussion of artistic media more generally.

schema: An agent's knowledge structure operating on various timescales and domains of information to make predictions about reality, including art events; analogous to literary-theoretic terms like *worldview*, *framework*, *mental model*, and *interpretive set*, and also to *PEM* terms like *top-down models* and *rules*.

sense-meaning: the intended signification of a text segment (word, phrase, string), which is situated within both the text's "coherence of senses" and the larger cultural context of the utterance.

surprisal: the degree to which data under a certain a occurrence is unlikely given the subject's models of reality (see Clark 2017).

WORKS CITED

Bakhtin, Mikhail. 2014. “1981[1934-1935]. ‘Discourse in the Novel’. In The Dialogic Imagination. Four Essays, 259–422, Selected 261–265, 268–275. Austin: University of Texas Press.” In *The Discourse Studies Reader*, 28–35.

Beaugrande, Robert de. 1982. “Text, Discourse, and Process: Toward a Multidisciplinary Science of Texts.” *Poetics Today* 3 (1): 195.

Beni, Majid D. 2017. “Commentary: The Predictive Processing Paradigm Has Roots in Kant.” *Frontiers in Systems Neuroscience*.

Bensman, Joseph, and Israel Gerver. 1958. “Art and the Mass Society.” *Social Problems* 6 (1): 4–10.

Bourdieu, Pierre. 1985. “The Market of Symbolic Goods.” *Poetics* 14 (1-2): 13–44.

———. 1998. *On Journalism and Television*. Pluto Press.

Bourdieu, Pierre, and Randal Johnson. 1993. *The Field of Cultural Production: Essays on Art and Literature*.

Bruner, Jerome S., and A. Leigh Minturn. 1955. “Perceptual Identification and Perceptual Organization.” *The Journal of General Psychology* 53 (1): 21–28.

Carhart-Harris, R. L., and K. J. Friston. 2019. “REBUS and the Anarchic Brain: Toward a Unified Model of the Brain Action of Psychedelics.” *Pharmacological Reviews*. https://doi.org/10.1124/pr.118.017160.

Clark, Andy. 2013. “Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science.” *The Behavioral and Brain Sciences* 36 (3): 181–204.

———. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, USA.

———. 2017. “A Nice Surprise? Predictive Processing and the Active Pursuit of Novelty.” *Phenomenology and the Cognitive Sciences* 17 (3): 521–34.

Clignet, Remi, and Diana Crane. 1988. “The Transformation of the Avant-Garde: The New York Art World, 1940-1985.” *Social Forces; a Scientific Medium of Social Study and Interpretation* 67 (1): 276.

Crane, Diana. 2015. “Avant-Garde Art and Artists.” In *International Encyclopedia of the Social & Behavioral Sciences*, 320–25.

Csikszentmihalyi, Mihaly, and Rick Emery Robinson. 1990. *The Art of Seeing: An Interpretation of the Aesthetic Encounter*. Getty Publications.

Davis, Murray S. 1971. “That’s Interesting!” *Philosophy of the Social Sciences* 1 (2): 309–44.

De Beaugrande, Robert, and Wolfgang U. Dressler. 2016. *Introduction to Text Linguistics*. Longman Linguistics Library.

Duchamp, Marcel. 1956. Marcel Duchamp interview on Art and Dada Interview by James Johnson Sweeney. Video. https://www.youtube.com/watch?v=DzwADsrOEJk.

Duckworth, Alistair M. 1979. “The Act of Reading: A Theory of Aesthetic Response. Wolfgang Iser.” *Nineteenth-Century Fiction* 34 (3): 337–43.

Dura, Marian T. 2007. “Sweet Anticipation: Music and the Psychology of Expectation. By David Huron. MIT Press, 55 Hayward St., Cambridge, MA 02142-1315; 800-405-1619; Http://mitpress.mit.edu. 2006. 512 Pp. Black-and-White Illustrations, Music Examples, Notes, Glossary, References, Index. Hardcover, $40.00.” *Music Educators Journal* 94 (2): 54–55.

Ekman, Ulrik. 2013. *Throughout: Art and Culture Emerging with Ubiquitous Computing*. MIT Press.

Evans, Owain. 2019. “Sensory Optimization: Neural Networks as a Model for Understanding and Creating Art,” November. https://arxiv.org/abs/1911.07068v1.

Friston, Karl J. 2013. “The Fantastic Organ.” *Brain: A Journal of Neurology* 136 (4): 1328–32.

Friston, Karl, James Kilner, and Lee Harrison. 2006. “A Free Energy Principle for the Brain.” *Journal of Physiology, Paris* 100 (1-3): 70–87.

Gjerdingen, Robert O., and Eugene Narmour. 1992. “The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model.” *Notes* 49 (2): 588.

Godfrey-Smith, Peter. 2016. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Straus, Giroux.

Gopnik, Alison. 1996. “The Scientist as Child.” *Philosophy of Science*. https://doi.org/10.1086/289970.

Gordon, Jan B., and Renato Poggioli. 1968. “The Theory of the Avant-Garde.” *The Journal of Aesthetics and Art Criticism* 27 (2): 244.

Grietzer, Peli. 2014. *Amerikkkka*. Vol. 004. GPDF Editions. Gauss PDF.

———. 2017. “A Theory of Vibe.” *Glass-Bead* 1 (1). https://www.glass-bead.org/article/a-theory-of-vibe/.

Griffiths, Thomas L., Charles Kemp, and Joshua B. Tenenbaum. n.d. “Bayesian Models of Cognition.” In *The Cambridge Handbook of Computational Psychology*, 59–100.

Hill, Grahame. 2001. *A Level Psychology Through Diagrams*.

Hohwy, Jakob. 2013. *The Predictive Mind*.

Hurley, Matthew M., Daniel Clement Dennett, and Reginald B. Adams. 2011. *Inside Jokes: Using Humor to Reverse-Engineer the Mind*. MIT Press.

Iser, Wolfgang. 1978. *The Act of Reading: A Theory of Aesthetic Response*.

———. 1993. *Prospecting: From Reader Response to Literary Anthropology*. JHU Press.

Kandel, Eric R. 2016. *Reductionism in Art and Brain Science: Bridging the Two Cultures*. Columbia University Press.

Kemp, Charles, Amy Perfors, and Joshua B. Tenenbaum. 2004. "Learning Domain Structures," June. http://dx.doi.org/.

Kemp, Charles, and Joshua B. Tenenbaum. 2008. "Structured Models of Semantic Cognition." *The Behavioral and Brain Sciences* 31 (06): 717.

Kesner, Ladislav. 2014. "The Predictive Mind and the Experience of Visual Art Work." *Frontiers in Psychology* 5 (December): 1417.

Koller, Daphne, Nir Friedman, and Francis Bach. 2009. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Levin, Phillis. 2001. *The Penguin Book of the Sonnet: 500 Years of a Classic Tradition in English*. Penguin Group USA.

Mandler, George. 2003. "Emotion." In *Handbook of Psychology, Research Methods in Psychology*, edited by Irving B. Weiner, John A. Schinka, and Wayne F. Velicer. John Wiley & Sons.

Martorella, Rosanne, and Diana Crane. 1988. "The Transformation of the Avant-Garde: The New York Art World, 1940-1985." *Contemporary Sociology* 17 (2): 221.

Narmour, Eugene. 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*.

———. 1992. *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. University of Chicago Press.

Nelson, Maggie. 2007. *Women, the New York School, and Other True Abstractions*. University of Iowa Press.

Noë, Alva. 2015. *Strange Tools: Art and Human Nature*. Hill and Wang.

Pinker, Steven, Martin A. Nowak, and James J. Lee. 2008. "The Logic of Indirect Speech." *Proceedings of the National Academy of Sciences of the United States of America* 105 (3): 833–38.

Rao, R. P., and D. H. Ballard. 1999. "Predictive Coding in the Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-Field Effects." *Nature Neuroscience* 2 (1): 79–87.

Reason, Suspended. 2016. "Every Little Star." Suspended Reason. May 31, 2016. https://suspendedreason.com/2016/05/31/every-little-star/.

———. 2017a. "Maps, Chords, and Effect Ideas." Suspended Reason. August 4, 2017. https://suspendedreason.com/2017/08/04/maps-chords-and-effect-ideas/.

———. 2017b. "Valencia/Rectify/Film/Television/Literature." Suspended Reason. October 2, 2017. https://suspendedreason.com/2017/10/02/valencia-rectify-film-tv-literature/.

———. 2018a. "On the Erotics of Interpretation." Suspended Reason. March 18, 2018. https://suspendedreason.com/2018/03/18/on-the-eroticcs-of-interpretation/.

———. 2018b. "Art as the Antithesis of Design." Suspended Reason. March 21, 2018. https://suspendedreason.com/2018/03/21/art-as-the-antithesis-to-design/.

———. 2018c. "Art vs. Design, a Follow-Up." Suspended Reason. March 23, 2018. https://suspendedreason.com/2018/03/23/art-vs-design-a-follow-up/.

———. 2018d. "Predictive Processing & Art as Cognitive Remodeling." Suspended Reason. March 27, 2018. https://suspendedreason.com/2018/03/27/predictive-processing-art-as-cognitive-remodeling/.

———. 2018e. "The Resonant." Suspended Reason. August 6, 2018. https://suspendedreason.com/2018/08/06/resonance/.

———. 2018f. "Schematic Disruption." Suspended Reason. August 17, 2018. https://suspendedreason.com/2018/08/17/schematic-disruption/.

Salimpoor, Valorie N., Mitchel Benovoy, Kevin Larcher, Alain Dagher, and Robert J. Zatorre. 2011. "Anatomically Distinct Dopamine Release during Anticipation and Experience of Peak Emotion to Music." *Nature Neuroscience* 14 (2): 257–62.

Salimpoor, Valorie N., David H. Zald, Robert J. Zatorre, Alain Dagher, and Anthony Randal McIntosh. 2015. "Predictions and the Brain: How Musical Sounds Become Rewarding." *Trends in Cognitive Sciences* 19 (2): 86–91.

Schmidhuber, Jürgen. 2009. "Driven by Compression Progress: A Simple Principle Explains Essential Aspects of Subjective Beauty, Novelty, Surprise, Interestingness, Attention, Curiosity, Creativity, Art, Science, Music, Jokes." In *Lecture Notes in Computer*

Science, 48–76.
———. 2010. “Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010).” *IEEE Transactions on Autonomous Mental Development* 2 (3): 230–47.
Schwarz, Wolfgang. 2013. “The Lure of Free Energy.” Wo’s Weblog. December 1, 2013. https://www.umsu.de/wo/2013/600.
Sharifian, F. &. Samani R. 1997. “Hierarchical Spreading of Activation.” In *Proc. of the Conference on Language, Cognition, and Interpretation*. IAU Press.
Stockwell, Peter. 2005. *Cognitive Poetics: An Introduction*. Routledge.
———. 2009. “The Cognitive Poetics of Literary Resonance.” *Language and Cognition* 1 (01): 25–44.
Swanson, Link R. 2016. “The Predictive Processing Paradigm Has Roots in Kant.” *Frontiers in Systems Neuroscience* 10 (October): 79.
Tenenbaum, J. B., C. Kemp, T. L. Griffiths, and N. D. Goodman. 2011. “How to Grow a Mind: Statistics, Structure, and Abstraction.” *Science* 331 (6022): 1279–85.
Turner, Fred, and Christine Larson. 2015. “Network Celebrity: Entrepreneurship and the New Public Intellectuals.” *Public Culture* 27 (1 75): 53–84.
Van de Cruys, Sander, and Johan Wagemans. 2011. “Putting Reward in Art: A Tentative Prediction Error Account of Visual Art.” *I-Perception* 2 (9): 1035–62.
Wampole, Christy. 2015. “On Awkwardness.” Guernica. June 18, 2015. https://www.guernicamag.com/christy-wampole-on-awkwardness/.