A Bayesian hermeneutic

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

Recently, cognitive scientists like Clark (2016) and Hohwy (2013), as well as computational neuroscientists such as Karl Friston (2006, 2013) have theorized the mind as a hierarchical prediction system, at levels varying from the "merely" sensory to the highly conceptual. Here, we extend this thesis by incorporating concepts from work in cognitive science on Bayesian knowledge structures. This synthetic model of a probabilistic, Bayesian or Bayesian-approximate mind is then employed as a means of understanding the hermeneutic process as it relates to textual and artistic encounters. We argue that one of the foundational mechanisms of the artwork, as it's contemporarily conceived, can be meaningfully conceptualized as an exploitation of the mind's predictive system. We further show how this mechanism, and a predictive cognitive framework, help explain a host of traditional literary, aesthetic, and art-historic 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*,
> cited in Gombrich 1960

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 (acquired knowledge). Though the problem of inference is "as old as recorded Western thought,"[1] we attempt here to sketch out a handful of precedents, especially relevant 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.[2,3] The schema anticipates predictive processing models by arguing that an organism's understanding of reality requires a synthesis of top-down models with bottom-up data.[1] In *Driven By Compression Progress* (2009), Schmidhuber notes a precedent for cognitive compression (a type of predictive process) in Piaget's theories of explorative learning and assimilation, where "new inputs [are] embedded in old schemas."[4]

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."[5] The

[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) and Beni (2017).[2,3]

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 it (*interpreter* function). Factors influencing perceptual set include expectations, emotion, 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 figures interpreted the middle figure as the letter "B."[6] Perceptual set was anticipated by Gestalt psychology, which argues the mind actively organizes incoming perceptions, through their interrelation, into gestalt wholes.

Around the same time, the emerging field of cybernetics had begun "framing purposive behavior" in systems of all kinds as "governed by feedback," self-motivatedly seeking out novelty with which to fuel said feedback system.[7] 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.[8] Gombrich's *Art and Illusion: A Study in the Psychology of Pictorial Representation* (1960), meanwhile, emphasized the role of schemata in both the production and reception of visual works.

In §1.2, we explain the fundamentals of predictive coding, prediction error minimization (PEM), and its speculated extension via concepts of Bayesian inference. Readers already familiar with these concepts may benefit from skipping ahead to §1.3 on hierarchical Bayesian models. Section 1.3, which focuses on the interfacing of PEM and Bayesian principles, will be especially of pertinence to those with research interest in high level predictive processing and the Bayesian brain. Section 1.4 provides a non-technical explanation of basic prediction and inference, availing itself most pertinently to non-technical readers.

Readers unfamiliar with predictive models of the mind, or with Bayesian inference, may benefit from consulting the GLOSSARY section at the conclusion of this paper.

1.2 From predictive coding to Bayesian error minimization

Predictive models of perception originated in computational neuroscience as a way of explaining low-level sensory processing such as in the visual and auditory fields [Rao and Ballard, as well as other low level PP work]. However, prominent cognitive philosophers, such as Andy Clark, and neuroscientists such as Karl Friston, have proposed that predictive processing and error minimization explains cognition at all levels of the brain and, through the free energy principle, is a fundamental property of life .[9–11]

Here, we will 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, including conceptual learning, and two, that the inference systems central to lower-level predictive coding are also 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), under the term *predictive error minimization*, Clark (2013/2017) under the handle *predictive processing*, and Friston (2013) as *free energy*.[9–12] As noted in Clark (2016), it has also been dubbed *active inference* in theories emphasizing its extension to motor control, and *hierarchical layered coding* by those emphasizing the organizational hierarchies believed to structure the different levels of prediction work.[13]

In this paper, we'll use the term prediction error minimization (PEM) as a generic or 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. Cast in probabilistic terms, the mind strives to learn models that approximate effectively (i.e. have minimal *Kullback-Leibler divergence* from) the probability distributions of experienced reality. 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."[5] 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 causes (e.g. high-dimensional data like social interactions and artworks), the hierarchical predictive structure must be able to encode high-level knowledge (i.e. store high-level patterns which explain and therefore compress multiple phenomena[2]) to make predictions.[9] Sets of predictions at the lower levels of the hierarchical structure are in turn anticipated at higher levels, cycling back. At the lowest level of the predictive structure, the system aims to predict sensory-data based on all higher levels. These predictions should provide a probabilistic evaluation of the sense-data sampled at each moment using knowledge represented across all layers.[3]

[2] Footnote on how PEM and Schmidhuber interact, or at least an interfacing/translation of terms of compression and prediction

[3] 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).

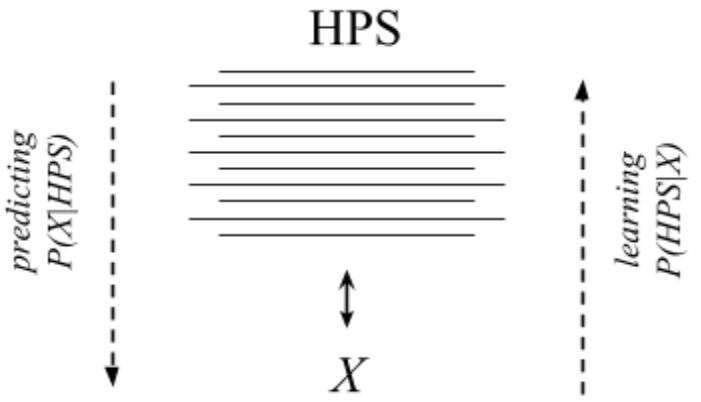

Fig. 1, showing the bidirectional nature of the hierarchical predictive system. HPS refers to hierarchical predictive structure; *X* is sensory data.

As described in Clark 2013 and 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 enable the brain to make predictions about, i.e. define a probability distribution over, sense data. Information represented at an arbitrary level of the 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 up increasingly 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 systems makes 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—used in order to make a particular prediction about sense data. The forward connections between layers direct a flow of concepts used to explain the observed stimulus, and backward connections enable updates to the structure and the adjusting of the priors relevant to understanding the stimulus.

One biologically plausible, computationally tractable[4] 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 (*umwelt*) experienced by the system. In Figure 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 | X). This distribution refers to the distribution over possible hierarchical predictive structures given observed sense-data X, and inference uses this quantity to updates the hierarchical predictive structure to best explain X[9,13]Bayesian inference we refer readers to [a, b, c].

We are interested in how predictive processing and the Bayesian framework can model a human observer's interaction with a piece 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 cognition related to art (even visual scene perception study is quite far removed...maybe explain briefly parenthetically), and the studies above use careful empirical research to justify their findings. 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 to aesthetic theory concerned with the interpretation of artwork.

1.3 HBMs as instantiations of hierarchical predictive structures

[4] Biological plausibility of Bayes principles as discussed in Clark 2013 section 2.2 and 3.3, Hohwy?, Friston? (see Clark 2013 sec. 3.3 for suggestions).

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 about 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[5] 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. If hierarchical predictive structures outline the flows of information to make predictions and learn from incorrect predictions, Bayesian principles provides us a broad 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" [Tenenbaum et al 2008, Kemp 2004]. The mind is not restricted to reasoning about one particular domain: it can utilize 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.

---

[5] 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 [Tenenbaum and Kemp].

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 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 The neighborhood example

An example of how, informally, inference works at higher conceptual and hermeneutic levels: Upon moving to a neighborhood of a new city, Italo begins taking morning walks to the bus station, often stopping for a coffee on his route. Because he comes into the office at different times depending on his workload, he observes the neighborhood at different times in morning.

Some morning Italo leaves for the bus stop before nine. In his old city, many shops opened early in the morning around seven or eight; his existing schemata for opening times reflect this and he behaves accordingly. In his new neighborhood, however, most shops appear to be closed so early in the morning, leading to a predictive failure which results in Italo regularly going coffee-less to work. Passively or actively, Italo has gathered bottom-up

information on when stores open on a given weekday and begins to make high-level inferences about the schedule of his neighborhood. .

At some tipping point of Bayesian modeling, Italo feel confident in making the induction that his new neighborhood doesn't open until nine. He tells this to his friend Alfonso, who mentions that from his own walks on Sunday he's noticed the neighborhood is usually closed that day until ten. Because Italo trusts Alfonso, he gives high credence to the information and incorporates it into his neighborhood schema. Italo then takes *action* on the knowledge, that is, he may start going into work later, or brewing his own coffee on mornings he wakes up early.

1.5 Schemata and predictive coding

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*.[14] Psychological and perceptual theories of schemata 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.[6]

We'll use "schema" or "schemata" 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 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.[12,13]

[6] 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.

In other more speculative contexts throughout this paper, “schema/ta” 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.

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.[8,15] Kandel cites modern art, for instance, as “dismantling perspective” in order to force the brain to “come up with a new logic of bottom-up processing.”[16] This paper departs by looking at artworks from a high-level, literary-theoretic lens (as opposed to a lower-level, VisArt perspective), surveying, first how hermeneutic processes rely on predictive processing, and second, how works of art and literature actively set up and subvert high-level thematic, narrative, and formal predictions by their audiences.

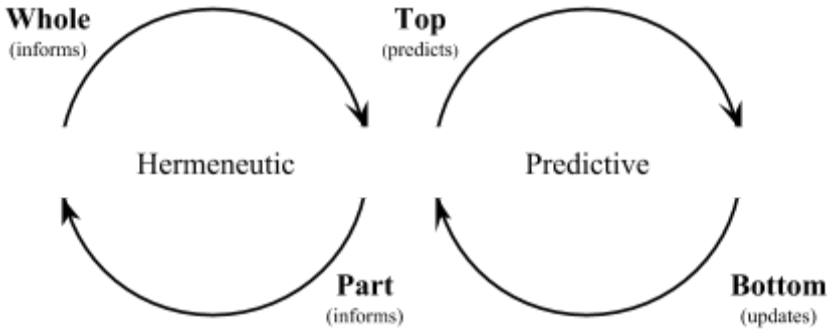

Fig. 2, depicting similarities between hermeneutic and PEM feedback loops in updating.

§2 Hermeneutics

All 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*” [emphasis his].[17] Elements are weighted by relevance and attention is allocated according to the configuration of the text and the schematic interests of the reader.

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.”[18]

2.1 Works as puzzles

All works of art, including texts, are puzzles insofar as they must be consciously or unconsciously “puzzled out” in order to yield coherence (in Bourdieuan terms, they must be “decoded” according to a soft grammar, which is to say, “parsed”). 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 or “decode” is already possessed by the majority of readers and can be applied with little friction. 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 requires the recognition of—to name but a few—sequential, logical, thematic, mythological, psychological, and ideological patterns within the work.[20] 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 Bayesian framework throughout §2.

2.2 Bayesian formulation of a hermeneutic encounter

In §2.1, we described the hermeneutic process as one in which "the reader or viewer is consistently but unconsciously interpreting the parts of the work, and how these parts might fit together in a coherent whole (see *Figure 2* above). Here we will formalize that process in Bayesian terms, representing a hermeneutic encounter graphically through an HBM, or hierarchical Bayes model.[19]

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 (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 engaged by 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. With this structure C and events X we can define a probabilistic joint distribution $p(X \mid C)$, where "joint" refers to a distribution of many random variables e.g. $\{x_1, x_2, \ldots, x_n\}$ and not merely one variable or event x. This distribution provides a probability of the events in X happening given knowledge structure C. A "coherent whole" might refer to a given structure over the sets X and C which yields a high value for $p(X \mid C)$, reflecting a high likelihood of the art events given our understanding of related concepts. While engaging with a work over time (e.g. reading sequential chapters of a book, interpreting elements of a painting one after another), we constantly make predictions about specific scenes and revise the possible outcomes for future remaining events 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, …, 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 of our knowledge structure C to understand

the next event in the piece, the remaining events of the piece, or the piece as a whole.[7]

In addition to making predictions about art events, we also learn from art events and update our knowledge structure 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[8]—a process markedly similar to that of the "hermeneutic circle" gestured at in *Figure 2*. This very well might happen piecewise during a work, where we calculate posteriors over subsets of concepts from subsets of the data *p(C' | X')*; a couple scenes might tell us about a few and not all of concepts relevant to a film. We can use algorithms from Bayesian networks, such as message passing, to do this type of inference on subsets as well as extrapolate to the entire structure. 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, a concept expanded on throughout §3.

Fleshing out the structure *C*: *C* is a graph (or Bayesian network, with a hierarchical structure resembling Figure 1 here). Each node of the graph can refer to a different concept or relationship between concepts at an arbitrary level of abstraction.. *C* contains all the concepts related to the work which might be deployed in other arenas (in the interpretation of other artworks, in understanding events in everyday life, in understanding history) as well as the structure which is solely concerned with relating the events *X* of a work.

By way of example, we can look at a few of the primary concepts, C, in play when watching a Spaghetti Western. One is a set of formal genre conventions that intersect with conventions of the era, based on other Westerns watched by the viewer; two, historical knowledge about the American West; three, general knowledge of human behavior put on display; and four, a general assessment of physical and operational possibility or "realism"—what we might call the

---

[7] 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.

[8] *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.

"soft physics" of the real and textual worlds. Crucially, these are merely the "central" or most 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*, 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 (see Appendix B, §1). Relevant concepts grow—but are also to an extent constrained, and altered—as the film carries out its runtime: there is sequential or "precedent" information, inference vis-a-vis causal structures, generalizations from relationships between parts, etc.

2.3 Interpretation and intentionality

Insofar as an observer sees a work 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 say*, and not merely what is, technically, said). Audiences search for markers of hierarchical importance in a communication, while artists employ established, widely understood cultural 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 (i.e. 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 loudness, prominence, size, position, or focus.

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. (Commas and syntax do not "mean" anything; they merely modify.) When the grammatical structure of a sentence is unclear, the interpreter must revert to probability assessments with respect to one, the real world; two, the probability of a statement's being possible or real; three, with respect to predictions about 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 deep learning and natural language processing). Is the announcement made “in trucks and vans”? Did the speaker announce the program in order to promote automobile safety, or is it the program that promotes automobile safety?[20]

In garden-path sentences, such as “The old man the boats,” the initial assumptions parse— that “old” was 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 inititial 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, must read a second or third reading necessarytime to verify that the hugely unlikely minority sense-meaning is, in fact, the one intended one here.

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 word within a given 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 have similar distribution models, this being the step that enables predictable communication, and thus 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, 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, and both interlocutors in the exchange share similar-enough models of the phrase’s rhetorical probabilistic likelihood within the threatened party’s linguistic schemata 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.

§3 Art and Bayesian inference

In §2, we considered how a hermeneutic encounter with an artwork could be framed in terms of prediction by an HBM. Specifically, we looked at the ways that unconsciously "puzzling out" (i.e. interpreting) a work allows for a coherence of meaning between parts. In much of avant-garde writing, however—as well as in specific literary traditions like modern (and post-) 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 even 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 nonetheless observe that there is a strong correlation between how "artsy" (how "fine," "refined," "literary," or "avant-garde") a work is perceived as in the cultural field, and the extent to which it actively resists, confounds, subverts, or problematizes the interpretive attempts and procedures of its audiences.[21,22] That is, the artsiness of art is a property defined by, or at least strongly correlating with, the work's confounding of, and resistance to, easy assimilation into observers' schemata.[9] Tellingly, art which does not complicate or subvert audience expectations is perceived as formulaic, predictable, or trite, and is accorded less prestige as a result.

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*

[9] 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—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 expectations of their audience.

about the world). Stockwell divides discursive modes into those which are *schema-preserving*, *schema-adding*, *schema-reinforcing*, and *schema-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.[23]

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 is therefore merely schema-preserving. *Second-order informativity* presents more "unusual" data and helps "develop schematic knowledge by accretion." *Third-order informativity*, finally, entails "highly unlikely" or unusual data with respect to the reader's status quo schema, and can represent such a disruption to said schema as to lead to its restructuring. For example, we might have a schema partway through a novel for "what kind of person" a literary character is. This schema can be represented by an HBM, resembling the one in §2, which encodes in its structure all knowledge useful to understanding the character and provides a probability distribution over the space of possible futures for the character. Said schema will require non-trivial revision should we come upon surprising new information about that character's actions or motivations later on in our experience of the text.[24]

These orders of informativity are well informed by the PEM framework. When the brain encounters high levels of "surprisal" in the world (i.e. encountering events that have low probability according to the mind's probabilistic understanding of the world i.e. the HGM) it may need to revise or restructure said probabilistic model that better explains the surprising data. The overlap between PEM and de Beaugrande's theories becomes clear, and we can understand the latter in terms of the former.

First-order information is low novelty and low surprise, resulting in the preservation of existing schema, or probabilistic models about the world. *Second-order information* is high novelty but low surprise, helping accrete and fill out existing schematic knowledge but stopping short of significantly altering any high-level probability models. Finally, *third-order information* is high novelty and high surprise, causing higher-level restructuring or significant

tuning of the reader's models. In first- and second-order information, some learning or inference in the HGM might occur, but the shift in the distribution after incorporating any error information from the original is less in magnitude than in third-order information

We'll refer to *forced hermeneutic revision* as a type of schematic subversion (schema-disrupting, in 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.

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 schemata. 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 experiences and learned information both within the space of the work and outside it. The sources or concepts, *C*, which inform expectations are, inevitably, one and the same with those that enable the act of interpretation in the first place(see §2.2).

In other words, effective art causes us to fall into predictive traps, into understandings or predictions about the artwork or the world of the artwork which are shown to be misguided.[25] Within the framework of cognitive poetics, art is schema-disrupting, and much of its value to audience members derives from third-order informativity. Tellingly, art that does not schema-subvert is deemed "predictable" or "obvious," and is regarded poorly as a consequence.

Schematic subversion can also be understood in the Murray Davis sense of *interestingness*.[26] Interesting information, to Davis, is that which denies the assumption-ground (set of assumptions) of its audience.

3.1.1 Bourdieu on "pure" poetry

> *"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.*[Pierre Bourdieu, 27]

3.1.2 Bowie's uniforms

> *People sometimes observe that David Bowie's personae drew attention to their own constructedness, and that's right in a sense. But really what Bowie's personae did was draw attention to their operational assumptions… [each] was the crisis of thinking its way out of its operational assumption.*[Peli Grietzer, 28]

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.[29]

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*" [emphasis ours].[29]

The philosopher Henri Bergson presents a similar analysis of the humor inherent in tripping and other physical or social awkwardnesses. Bergson argues that "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."[30]

When one's predictions fail in an art encounter, it is much like Bergson's jolt: up and out of the automaticity of interpretation (as well as the schema responsible for the misinterpretation in the first place). Because action is required to correct the prediction error, cognitive *arousal* is induced.

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."[23]

Readers and viewers therefore encounter "charged" scenarios—data points which because they represent rich, complex phenomena in the real world are highly meaningful, and which are 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 worldview which is itself tethered to the world.

In §4.5, 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.

If in humanistic, cartographic art, schematic subversion is typically of a charged subject typically relates to the real world, for instance social psychology. In more formally oriented works, the subverted subject might relate to specifically linguistic or graphic expectations, such as the probabilistic sense-meanings of a word (§2.3) or the unspoken rules and boundaries of a discipline. Artists like Duchamp and Cage epitomize the latter approach, pure

poetry are chock-full with the former, though it is a foundational priority of modernist and twentieth century art practices more generally.

3.3.1 Ethical modeling and subversion in Whit Stillman's *Metropolitan*

Nick Smith: You know what kind of guy [Rick von Slonecker] is.
Cynthia: I don't, it's just what you say, and you're completely untrustworthy.
Rick von Slonecker: Smith, you're a liar. I've heard the crap you've been telling about me.
Nick: Aw, you have?
Rick: Yeah about some girl I supposedly mistreated, Polly Perkins? There isn't any Polly Perkins, you know that. Such a girl never even existed. Go on, tell them, you made it up.
Cynthia: Is this true?
Nick: Yes and no.
Charlie: Oh god, Nick.
Cynthia: You did make it up.
Nick: There is no one Polly Perkin. There are many of them.
Rick: So you admit you lied!
Nick: Polly Perkins is a composite, like *New York Magazine* does.
Rick: Name one girl.
Nick: Girls that have been degraded by you don't need the further humiliation of having their named bandied about non-exclusive afterparties.
Rick: You see, there isn't one.
Charlie: This looks really bad, Nick.
Nick: Cathy Livingston. No more harm can be done to her now.
Rick: I had nothing to do with that, and anyway she was completely unstable.
Nick: That didn't stop you from boozing her up and then talking her into pulling a train for you and Lenley!
Rick: That's not how it happened.
Nick: Aw, what did happen, Rick? [...] She tried to call and talk to you and you wouldn't even see her… Are her panties still in your collection? Rick keeps a collection of the girls he seduced. When they kill themselves do you do anything special to memorialize them?

3.3.2 Examples of poetic subversion from recent readings

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 and that are revised by "Ovaltine." In the second, "too" is probabilistically modeled as meaning "as well," an assumption which is revised 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 nakedness. Mayer's writing has been praised for its subversion of reader understanding: "[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*.[31]

3.4 Hermeneutic revision and the effect idea[32]

Schematic subversion is an effective way of not just adjusting but also of baring expectations (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. Described as 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. 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 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 schema, thus we can classify the effect idea, in accordance with Stockwell's modes of discourse, as *schema-baring*.[10]

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. Peli Grietzer, 2014: "Art is more 'artsy' the more [the hermeneutic] process bolsters its intended impact."[33] Tellingly, conceptual art hinges on the effect idea, using the subjected experience as a way to communicate—a mode both less direct and less "railroad-y" than alternative models.

3.4.1 Spielberg's subversion

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 audiences expectations (that the film's subject will be extraterrestrials) by showing two bright beams 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 (trailer-cued 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 "other" is revealed in fact to be an instrument of man, an effect which carries within it an implicit set of ideas.

[10] To be clear, 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* and *schema-baring* functions of art objects within the predictive framework introduced in §1 and 2. Here, in §4 and 5, we speculate possible attributes of art's interactions with audience schemata.

4.1 Art as superstimulus

Following Hurley et al. (2011), we propose that art is, among other functions, a kind of higher-level cognitive superstimulus culturally evolved to exploit 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 walk listeners down a specific interpretation of events only to pull the rug out,[29] art is tailored to target existing compressions in a subject's (parts of which being often culturally shared) schemata. What follows in an intense encounter between a subject and his schema on one side and the art object with its highly compression-prone information on the other.

4.2 Self-supervision and truthiness

In an art encounter our mind updates its inferential models about the world based on causal relationships and concepts it believes to be true, rather than those which actually govern the reality's workings (the latter condition is not available to a observer). In this way, we can, borrowing a machine learning term, conceptualize the mind as *self-supervising*. A consequence of self-supervision is that information which fits closely with existing schemata, but poorly with reality, is perceived as more, rather than less, likely to be the case.

The observer's schema acts as a "check" or arbitrator on its own incorporation of the art works' worldview (dynamics and concepts);[11] when the cartography

[11] 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.

of the work is too implausible in the eye's of an apprehending schema, it is dismissed entirely. This is to say that, given our standing understandings of the world, concepts learned directly (that is from personal experience) and indirectly (that is, from outside information sources) are used to assess the likelihood of a work's worldview as conveyed through its components. Following cognitive scientist Endel Tulving, we can refer to personally experienced concepts as *episodic*, and to second-hand and scientific knowledge as *semantic*, together constituting the observer's schema. In this way the schema can be understood to work as a gatekeeper to its own revision.

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,[17] and the vagueness of the term as used in non-scholarly discourse fields has led to its being avoided in academic discourse.

As we will use it here, to resonate is to oscillate in such a way that "a sympathetic oscillation occurs in a similar nearby structure,"[34] 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, their worldview, their existing psychic schemata of the world.

Stockwell does the valuable work of researching the physics of resonance in order to elaborate on its metaphor. From him, we learn, resonance 1) is measured in *intensity*, 2) is reduced by *dampening effects*, 3) has a *decay*, the length of time it takes for a reverberation to revert to silence, and 4) can transform into *echo* when the dampening effect is inadequate, causing a positive feedback loop between the original oscillating body and the resonating body which has picked up its vibrations.[17]

While resonance is one of our key checks in self-supervised art learning, it is prone 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.[34]

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, the analogous "credibility" is similarly important as a marker of the accuracy of presented information, or the efficacy of an action proposal.

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.[35] This arousal is neutral in that the situational context of its occurrence influences whether it appears to the subject as a positive emotion (e.g. interest) or a negative one (e.g. stress).[15]

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 "important changes in the environment" of the work. On some level, even as much as one "mistakenly" views an artwork as real-world data to populate its models, the artwork is still cordoned off as a sense-datum of minimal consequence, a zone of low-risk and therefore low-anxiety learning. 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 pleasurable rather than stressful.[25]

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.[15][8] Phrasing this in the framework of the paper, we might say that an observer enters an art encounter with a set of interpretive concepts, *C*, appropriate to the work. When those concepts (and therefore the larger schema) are defied (as they inevitably are, in

non-superfluous, meaningful works), the observer must search for and/or build alternative models by which to understand the work, and the eventual discovery of an (not “the”) appropriate schema produces subjective pleasure.

This model finds support in Schmidhuber and Friston, who both ascribe pleasurability to efficient, instrumentally accurate explanations (or “compressions”) of previously unpredictable stimuli.[36–38]

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) in which the subject's attention is focused primarily on the stimulus at hand. Attention, as Clark notes, 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.[13] After the fact, art encounters are frequently contemplated at length, often socially or textually (e.g. critical reception) in discursive scenarios not dissimilar to Hohwy’s model of *introspection as inference* found in 2013’s *The Predictive Mind*.[12]

Further, the concept of *precision* captures the estimated “inverse variance of a prediction error signal—in other words, it sets error bars around an error signal according to its currently estimated importance or reliability.” “High-precision errors enjoy greater post-synaptic gain and (hence) increased influence,”[11] and we can easily 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 specifically 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 cast a handful of these concepts

in a predictive framework, hopefully illustrating PEM's high 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 Csikzentmihalyi & Robinson 1990) "on the outside, unable to interact with the work."[39] In an environment dominated by (relative) 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 about the work. This phenomenon might reflect the activation by the artwork of a meager amount of relevant concepts (i.e. nodes) in the viewer's knowledge structure.

5.1.1 Domesticity

Related to accessibility is the concept of *domesticity* as defined by Thurston 2018. When a textual work interfaces with many concepts, *C*, familiar and highly networked in the mind of reader *R*, there will be a high proportion *spreading activation*, the process by which textual materials contact and interface with associated concepts within *R*'s mind. We can call such works "domestic" for their high relevance and comfortableness to an observer.[40]

5.1.2 Genericism

When the information and concepts necessary to access a work are common to large swathes of the population (or when a work's thematic concepts can be considered widely domestic), we can call this work both highly accessible and highly *generic*.[cf. 41]

5.2 Ideal and implied readers

While 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 points out) would find the author's text entirely "superfluous,"[42] 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."[42] 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 (assumes both conceptually and also with respect to the probabilistic "sense-meanings" in the language, *cf.* §2.3 & glossary) and subverts in his writing of the text.

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 include time and space (e.g., following Clark, "that there [can] only [be] one object... in one place, at a given scale, at a given moment"). Hyperpriors, in our Bayesian formulation of art encounters, are high-level priors that constrain the space of possible subnetworks of concepts and their relations an observer formulates from experiences and art encounters and a dominant hypothesis selected.[2,9] As such, all hermeneutic work is guided or biased by them, and they tend towards "fixed" as opposed to "fungible"—for instance, hyperpriors are less up for grabs", or susceptible to revision, in an art encounter than hypotheses regarding the work's meaning or content (though many artists do, in fact, attempt to misguide audiences with faked or absent hyperpriors, such as by keeping their identities mysterious through pseudonyms).

Relevant hyperpriors in an art encounter include: *knowledge of the artist, knowledge of the artist's reputation, knowledge of the work's reputation, knowledge of the work's market value, knowledge of the artist's biography, knowledge of the temporal and(sub)cultural habitus, i.e. chronotopic frame, into which the work was produced,* etc.[cf. 11 §1.4]

5.4 Defamiliarization

A well-known phenomenon in psychology is the cessation of full awareness of familiar stimuli. The phenomenon has been known as "compiling" by Herbert A. Simon, "tacit dimensionality" by philosopher Michael Polanyi, and the "ready-to-hand" by Heidegger.[43] 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.

To the Russian formalists, a crucial aim of art was defamiliarization, where the everyday and banal (or "ready-to-hand," or "already 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 familiar-known, an analogic link is created between the two which upcycles into new models of the familiar.[12]

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."[29] Tellingly, both awkward situations (*cf.* §3.2) and artistic encounters are characterised by the disruption of automaticity in favor of more disfluent hermeneutic or affective states.

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 schemata occupy a high enough probability of possibility 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 (as with 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

[12] 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.'"

racial regressiveness in the segregated era. The 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 his presentation of an object or occurrence so that it is unclear whether attention ought be paid—e.g. by placing signal in a 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 sort of meta-distribution) has been altered, which will change the structure of future predictions.[20]

5.6 Predictive coding and music

Music is the arts discipline which has been most interfaced 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 rewarding to listeners, whereas more familiar music, in which listeners can actively anticipate coming notes and chord shifts, is more rewarding.[44–46] These observations have been 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.[47]

In musicology, a relative of predictive processing models has emerged in Eugene Narmour's implication-realization model of melodic expectation.[48–50]

5.7 Structural "schemata" and patterned influence

> Art is the imposing of a pattern on experience and our aesthetic enjoyment is recognition of the pattern.
>
> Alfred Whitehead, *Dialogues*

One of the primary ways artists learn their craft is through repeated, thoughtful exposure to existing works in their discipline (as well as analogically, by

observation of works from other disciplines, in a *radial proximity* fashion). If, as theories of hierarchical predictive error minimization and processing have claimed, actions and motor control are tied up in perceptual patterns and top-down predictions, work may be possible to understand structural "schemata" (of the type theorized by musicologist Robert Gjerdingen's in his touchstone *Music in the Galant Style[51]*) and patterns of influence (or influences of patterns) between consuming and creating artworks. These structural patterns can be understood as recognizable, built-in chunks that lend familiarity and pleasurable predictability to unfamiliar works, since they have been copy-paste transmitted across the field of production — a mechanism for prediction error minimization similar to the conception of "culture as patterned practices"[52] outlined in §5.8 on mass culture.

5.8 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 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.[9]

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 mass media like film creates more homogeneous (and thus predictable) cultural rituals, for example in romantic norms. Thus the literary-theoretic concepts of texts as "models" (such as by Iser 1993)[53] takes on further meaning.

5.9 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 the 19th and 20th centuries.[22] Insofar as the subversion, confounding, resisting, and problematizing of interpretation can be understood as purposeful illegibility (at least to an extent), the world of design is characterized by its adherence to the principle of supreme legibility. Successful design appears to require insight into audience (or "user") schemata 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). [21,22,54]

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.[22]

5.9.1 Temporal decay in consideration as "art"

Moreover, as art 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[55]) hyper-specific and thus no longer in play for contemporary audiences. 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.[22]

5.10 Future directions?

This paper cannot provide a full interfacing of literary theoretic, art historical, and aesthetic frameworks with theories of predictive processing and error minimization. We'll close only by laying out a handful of concepts which jump readily to mind as potentially congenial with theories of PEM: *critical reagents*, *free indirect discourse*, *gappiness*, *genre*, *indirect speech act*s, *interestingness (as defined by Murray Davis, "That's Interesting!")*, *irony*, *the logical consistency of fantasy worlds*, *punning*, *red herrings*, *representation*, *rhetoric*, *rhyme*, *suspense*, *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 (calculating the probability of an event given prior knowledge) is applied to known information in order to update a model's priors toward accuracy

bidirectional: referring to the way, in multilevel predictive systems, bottom-up information influences top-down *schemata* and vice-versa

concept, C: a hierarchical knowledge structure consisting of concepts and relations between concepts which are engaged by a set of events or an engagement.

cartographic function: the quality of an artwork "mapping" or modeling reality

compressability: regularity of a data set, either inherent or with respect to some pattern already in the subject's *knowledge structure*

compression: the distilling of data, via noticed regularities, into *simplifying explanations*

curiosity: desire to create or discover more compressible data, often toward some overall compressive goal or "progress" (*cf* Schmidhuber, Clark)

gain: the "volume" or "urgency" of the error signal, influenced in part by the incoming information's *precision*

hermeneutics: the practice of interpreting texts

hierarchical Bayes model: a class of Bayesian models which learn and reason probabilistically about reality by employing knowledge encoded in the models' structures, making predctions and learning from prediction failures.

hierarchical generative model/s (HGM): Clark's specific carving of the predictive *schema/ta*

hyperpriors: a systemic prior, tending toward the abstract and fundamental, in subjects' predictive models (*cf* Hohwy, Clark)

interestingness: a subject-dependent phenomenon triggered when an observer learns to predict or compress said data in a more efficient way (*cf* Schmidhuber, Davis)

knowledge structure: the set of concepts and their relationships constituting an HGM and guide its predictions and probabilistic evaluations of the environment, i.e. allow it to make probabilistic evalautions of

Kullback-Leibler divergence: a statistical measurement of *prediction error*

mutual prediction error: a theory proposed by Clark that social and cultural systems are built in part to minimize mutual prediction error by establishing common patterns of behavior

noise: small predictive errors that can be ignored, incoming information left out of models as overly random (low *compressability*) or irrelevant (and thus excluded from conscious attention)

precision: the reliability of information in an inferential system, analogous to a signal-to-noise ratio

prediction error: the distance between the expected (i.e. modeled) and the actual.

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

PEM: prediction error minimization, a cognitive science theory linking the *predictive coding* of computational neuroscience with cog sci's *hierarchical generative models* and *Bayesian inference*

observer: term encompassing both VisArt and film's "viewer," and literature's "reader," allowing for a liberal discussion of artistic media at once

schema/ta: adapted from psychology; for our purposes includes the subject's Bayesian *priors* or *probability distributions* of expectations vis-a-vis reality as well as the hierarchic quality of Clark's multilevel systems… in its plural form *schemata*, denoting a full or subset of the subject's models, is analogous to literary-theoretic terms like *worldview*, *framework*, *mental model*, and *interpretive set*, and also to *PEM* terms like *top-down models* and *rules*.

simplifying explanations: the result of compressing regularities in a data set down to their common denominators/underlying algorithms

surprisal: data whose occurrence is unlikely given the subject's models of reality (*cf* Clark)

APPENDIX A: SECTION REFERENCE

1.3 Lakoff analogy as cross-domain sharing of priors

Thus far we have been discussing hierarchical inference, but cross-domain mapping allows horizontal, analogic inference as well. As Lakoff (1992) notes, much of the language we use to talk about abstract domains is merely a figurative extension of language used to discuss analogues physical-world situations. For example, in the analogy *love is a journey*, physical journeys, in which a subject crosses a physical distance over a period of time, are used as a metaphor for romantic love (*a dead-end street; a long and bumpy road; can't turn back now; at a crossroads; spinning the wheels; off-track; on the rocks; taking control of the whee; bailing out*).

More importantly for our purposes, Lakoff further argues that we reason about abstract and metaphysical phenomena specifically through inferential work regarding analogous physical phenomena. Thus inferential patterns used to reason about travel and journeys are often used in reasoning about romantic relationships; the solutions to getting figuratively "stuck" are analogized versions of the solutions to getting literally stuck (e.g. in the mud, etc). Further, it is likely this process happens according to the *radial proximity* of adjacent situations, that is,

weakly analogous cases will have less influence than strongly analogous cases.[13]

At the Bayesian level, analogic can be theorized as the search for some connection between distinct high-level concepts. Expressed in Bayesian terms, analogy might be like adding a connection between nodes in B, a subset of C, which contains nodes at a high-level in the knowledge structure consisting of concepts C over some arbitrary graph structure. We make no constraints on the class of domains and timescales for B. The new connections change each node's set of evidence and priors as well as all the nodes in B which might be involved the update step of an inference algorithm. High-level knowledge comprised of different domains and concepts have now been put in contact.

In the case of predictive error minimization and processing, we can imagine that the analogous situations might engage a common set of distributions, or *priors*, especially when one knowledge domain is especially weak and thus predictions high-error. In asymmetric situations such as Lakoff's *death as night* (the physical-literal used for inference in analogous abstract-figurative situations), however, it seems intuitively likely that the inferential influence is similarly asymmetrical, that is, a subject's priorsand inferential work regarding e.g. night would be affected by new subjective understandings of death.

2.2-2.4, 5.8 Conversational co-construction

Meant as accompaniment for §2.2-2.4, especially w/r/t the reader-response-esque "co-construction of a conversation." Also in conversation with §5.8 on mass cultural and mutual error prediction via established patterns of social behavior:

As "explored in some detail by Pickering and Garrod (2007, 2013) is the co-construction of a conversation. In conversation, Pickering and Garrod suggest, each person uses their own language production system (hence the generative model underlying their own behaviour) to help predict the other's

[13] See also similarity-based theories of inference, as well as Kemp and Tenenbaum 2009, which speculates four models of inductive work through reference to prior knowledge (for our purposes, "analogy"). Those are: taxonomic models (i.e. relatedness), spatial models (analogy to Euclidean space), threshold models (delimited gradients), and causal models (if-then).[56]

utterances, while also using the output of the other as a kind of external scaffolding for their own ongoing productions. These predictions (just as PP would suggest) are probabilistic, and span multiple different levels from phonology to syntax and semantics."[13]

Psychologist and archaeologist Duarte and Stefanakis (2014), respectively, stumble upon a related idea — sans predictive mechanisms: a work of art "must 'catch' our attention by activating a specific significant reaction it it. It has even been defended that art appreciation is motivated by the pleasure provoked by attention arousing... Now, our nervous system tends to be particularly activated by artistic representations, which not only reproduce the essence of the objects symbolized but actually amplify them, on the form of 'super-stimulus'"[57 trans.]

WORKS CITED

1. Griffiths TL, Kemp C, Tenenbaum JB. Bayesian Models of Cognition. The Cambridge Handbook of Computational Psychology. pp. 59–100.

2. Swanson LR. The Predictive Processing Paradigm Has Roots in Kant. Front Syst Neurosci. 2016;10: 79.

3. Beni MD. Commentary: The Predictive Processing Paradigm Has Roots in Kant. Front Syst Neurosci. 2017;11: 98.

4. Schmidhuber J. Driven by Compression Progress: A Simple Principle Explains Essential Aspects of Subjective Beauty, Novelty, Surprise, Interestingness, Attention, Curiosity, Creativity, Art, Science, Music, Jokes. Lecture Notes in Computer Science. 2009. pp. 48–76.

5. Hill G. A Level Psychology Through Diagrams. 2001.

6. Bruner JS, Leigh Minturn A. Perceptual Identification and Perceptual Organization. J Gen Psychol. 1955;53: 21–28.

7. Turner F, Larson C. Network Celebrity: Entrepreneurship and the New Public Intellectuals. Public Culture. 2015;27: 53–84.

8. Kesner L. The predictive mind and the experience of visual art work. Front Psychol. 2014;5: 1417.

9. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 2013;36: 181–204.

10. Friston KJ. The fantastic organ. Brain. 2013;136: 1328–1332.

11. Clark A. A nice surprise? Predictive processing and the active pursuit of novelty. Phenomenol Cognitive Sci. 2017;17: 521–534.

12. Hohwy J. The Predictive Mind. 2013.

13. Clark A. Surfing Uncertainty: Prediction, Action, and the Embodied Mind. Oxford University Press, USA; 2016.

14. Reason S. Schematic Disruption. In: Suspended Reason [Internet]. 17 Aug 2018 [cited 26 Aug 2018]. Available: https://suspendedreason.com/2018/08/17/schematic-disruption/

15. Van de Cruys S, Wagemans J. Putting reward in art: A tentative prediction error account of visual art. Iperception. 2011;2: 1035–1062.

16. Kandel ER. Reductionism in Art and Brain Science: Bridging the Two Cultures.

Columbia University Press; 2016.

17. Stockwell P. The cognitive poetics of literary resonance. Lang Cogn. 2009;1: 25–44.

18. De Beaugrande R, Dressler WU. Introduction to Text Linguistics. Longman Linguistics Library; 2016.

19. Sharifian F&. SR. Hierarchical spreading of activation. Proc of the Conference on Language, Cognition, and Interpretation. IAU Press; 1997.

20. suspendedreason. Valencia/Rectify/Film/Television/Literature. In: Suspended Reason [Internet]. 2 Oct 2017 [cited 27 Aug 2018]. Available: https://suspendedreason.com/2017/10/02/valencia-rectify-film-tv-literature/

21. Reason S. Art as the Antithesis of Design. In: Suspended Reason [Internet]. 21 Mar 2018 [cited 24 Oct 2018]. Available: https://suspendedreason.com/2018/03/21/art-as-the-antithesis-to-design/

22. Noë A. Strange Tools: Art and Human Nature. Hill and Wang; 2015.

23. Stockwell P. Cognitive Poetics: An Introduction. Routledge; 2005.

24. de Beaugrande R. Text, Discourse, and Process: Toward a Multidisciplinary Science of Texts. Poetics Today. 1982;3: 195.

25. Reason S. Predictive Processing & Art as Cognitive Remodeling. In: Suspended Reason [Internet]. 27 Mar 2018 [cited 26 Aug 2018]. Available: https://suspendedreason.com/2018/03/27/predictive-processing-art-as-cognitive-remodeling/

26. Davis MS. That's Interesting! Philos Soc Sci. 1971;1: 309–344.

27. Bourdieu P. The market of symbolic goods. Poetics . 1985;14: 13–44.

28. Grietzer P. Peli Grietzer on Twitter. In: Twitter [Internet]. 27 Aug 2018 [cited 27 Aug 2018]. Available: https://twitter.com/peligrietzer/status/1034077959030820864

29. Hurley MM, Dennett DC, Adams RB. Inside Jokes: Using Humor to Reverse-engineer the Mind. MIT Press; 2011.

30. Wampole C. On Awkwardness. In: Guernica [Internet]. 18 Jun 2015 [cited 26 Aug 2018]. Available: https://www.guernicamag.com/christy-wampole-on-awkwardness/

31. Nelson M. Women, the New York School, and Other True Abstractions.

University of Iowa Press; 2007.

32. Reason S. Maps, Chords, and Effect Ideas. In: Suspended Reason [Internet]. 4 Aug 2017 [cited 26 Aug 2018]. Available: https://suspendedreason.com/2017/08/04/maps-chords-and-effect-ideas/

33. Grietzer P. Amerikkkka. Gauss PDF; 2014.

34. Reason S. The Resonant. In: Suspended Reason [Internet]. 6 Aug 2018 [cited 26 Aug 2018]. Available: https://suspendedreason.com/2018/08/06/resonance/

35. Mandler G. Emotion. In: Weiner IB, Schinka JA, Velicer WF, editors. Handbook of Psychology, Research Methods in Psychology. John Wiley & Sons; 2003.

36. Friston K, Kilner J, Harrison L. A free energy principle for the brain. J Physiol Paris. 2006;100: 70–87.

37. Schmidhuber J. Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010). IEEE Trans Auton Ment Dev. 2010;2: 230–247.

38. Schwarz W. The lure of free energy. In: Wo's Weblog [Internet]. 1 Dec 2013 [cited 26 Aug 2018]. Available: https://www.umsu.de/wo/2013/600

39. Csikszentmihalyi M, Robinson RE. The Art of Seeing: An Interpretation of the Aesthetic Encounter. Getty Publications; 1990.

40. Thurston H. Narrative Domestication. In: The Sublemon [Internet]. 17 Aug 2017 [cited 31 Aug 2018]. Available: http://thesublemon.tumblr.com/post/164283154217/narrative-domestication

41. Reason S. Generic Fit. In: Suspended Reason [Internet]. 8 Jul 2016 [cited 26 Aug 2018]. Available: https://suspendedreason.com/2016/07/08/generic-fit/

42. Iser W. The Act of Reading: A Theory of Aesthetic Response. 1978.

43. Ekman U. Throughout: Art and Culture Emerging with Ubiquitous Computing. MIT Press; 2013.

44. Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat Neurosci. 2011;14: 257–262.

45. Salimpoor VN, Zald DH, Zatorre RJ, Dagher A, McIntosh AR. Predictions and the brain: how musical sounds become rewarding. Trends Cogn Sci. 2015;19: 86–91.

46. Dura MT. ■ 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. 2007;94: 54–55.

47. Reason S. Every Little Star. In: Suspended Reason [Internet]. 31 May 2016 [cited 27 Aug 2018]. Available: https://suspendedreason.com/2016/05/31/every-little-star/

48. Gjerdingen RO, Narmour E. The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model. Notes. 1992;49: 588.

49. Narmour E. The Analysis and Cognition of Basic Melodic Structures: The Implication-realization Model. 1990.

50. Narmour E. The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model. University of Chicago Press; 1992.

51. Gjerdingen R, Bourne J. Schema Theory as a Construction Grammar. Music Theory Online. 2015;21. doi:10.30535/mto.21.2.3

52. Roepstorff A, Niewöhner J, Beck S. Enculturing brains through patterned practices. Neural Netw. 2010;23: 1051–1059.

53. Iser W. Prospecting: From Reader Response to Literary Anthropology. JHU Press; 1993.

54. suspendedreason. Art vs. Design, a follow-up. In: Suspended Reason [Internet]. 23 Mar 2018 [cited 24 Oct 2018]. Available: https://suspendedreason.com/2018/03/23/art-vs-design-a-follow-up/

55. Bakhtin M. 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. The Discourse Studies Reader. 2014. pp. 28–35.

56. Kemp C, Tenenbaum JB. Structured statistical models of inductive reasoning. Psychol Rev. 2009;116: 20–58.

57. Duarte AM, Stefanakis MI. El uso de señales para la atención en el arte Griego antiguo: aspectos que influencian la concentración en la obra de arte y sus elementos. Arte, Individuo y Sociedad. 2015;27. doi:10.5209/rev_aris.2015.v27.n3.46654