Art, meaning, and aesthetics: the case for a cognitive neuroscience of art

William P. Seeley

Bates College

Author's Note: This manuscript is a draft of a chapter to appear in eds. Marcos Nadal, Joseph P. Huston, Luigi Agnati, Francisco Mora, and Camilo José Cela-Conde, Art, Aesthetics, and the Brain (New York: Oxford University Press). Please address all communication to Bill Seeley at wseeley@bates.edu. Please do not cite without permission of the author.
Art, meaning, and aesthetics: the case for a cognitive neuroscience of art

It is important to note that I am not suggesting that we should directly import the results of empirical psychology to aesthetics. The direct application of empirical results in aesthetics can, and very often does, go terribly wrong. What I suggest is that aesthetics should take some new paradigms of philosophy of perception seriously. The specific paradigm I am interested in here, the paradigm of multimodality, is based on a large body of empirical research. However, my aim is not to urge an empirical turn in aesthetics, but to urge a turn in aesthetics towards philosophy of perception, and this sometimes entails a turn towards empirically informed philosophy of perception (Nanay, 2012, 353-354).

The philosophy and psychology of art often behave as estranged bedfellows, fighting tooth and nail over absent common ground that neither concedes could exist. Empirical aesthetics is predicated on the claim that philosophical studies of art are methodologically suspect. The latter field has been defined as a *speculative aesthetics* grounded in deductive inferences drawn from the subjective judgments of experts. On this account philosophy of art, at best, reflects the subjective tastes of critics and theorists (Berlyne, 1971; Martindale, 1984). Philosophers behave no better. They claim that it is a category mistake, an error in logic, to ground a theory of art in behavioral studies -- questions about the nature of art are questions about the structure and validity of artistic conventions, questions that can be no more answered by studying the preferences and physiological responses of large groups of naïve subjects than analogous questions about the validity of moral beliefs (Dickie, 1962; Wittgenstein, 1966; see also Carroll, Moore, and Seeley, 2012; Seeley and Carroll, 2013). My suggestion, borrowing the letter, but perhaps not the intent, of Bence Nanay's comment above, is that there is something to be learned in a methodological rapprochement on both sides of this great divide. Philosophical questions about the nature of artistic salience and the role evaluative judgments play in our interaction with artworks can be modeled as questions about the influence of cognitive and affective processing in perception. Philosophers would
do well to pay attention to psychological theories about how these variables influence our engagement with artworks. Likewise, psychology of art attributes a role to aesthetic appraisals in our engagement with artworks that is not consistent with the way artistic salience is attributed to artworks in contemporary artistic practices. Empirical aesthetics would, therefore, do well to pay attention to the range of ways way philosophers of art model cognitive, affective, and aesthetic influences on judgments of artistic salience. In what follows I propose a model for a cognitive neuroscience of art to bridge the gap between philosophy of art and empirical aesthetics and discuss how the model resolves two standard philosophical objections to empirical aesthetics.

1. Art, Aesthetics, and Psychology

The history of mistrust between philosophy of art and empirical aesthetics can be traced to a disagreement about the nature of artistic salience. Think of it this way. Why would the judgments and intuitions of art critical experts be of any interest at all to philosophers of art? This seems like odd evidence out of which to build an objective, general theory of art! Well. Ontological questions about art are to a large extent questions about the conventions governing our categorical (and evaluative, but more on that later) judgments about a range of artifacts. If we want to know about these conventions, who should we ask -- or better, who's behavior should we be interested in -- experts or non-experts? It seems like the place to start is with people who understand the language so to speak (Wittgenstein, 1958; Dickie, 1962). The ability to evaluate the judgments of non-experts in these contexts depends on a prior understanding of the structure of these conventions, conventions that are embedded in the behaviors of art experts. I suppose the worry expressed by empirical aesthetics is that this might render artistic judgments subjectively arbitrary and elitist. However, this doesn't necessarily follow. The claim is not that the judgments of experts arbitrarily determine the artistic conventions of a community. It is rather that expert judgments codify the shared practices that define an artistic community and in so doing reflect the conventions against which artistic judgments and behaviors are evaluated within that community.

Research in empirical aesthetics replaces the explicit judgments of expert consumers with
physiological measures of arousal constitutive of (at least) the feeling (and in some theoretical contexts the content as well) of aesthetic experience. These physiological measures are treated as objective measures of *aesthetic interest and appraisal* that generalize across art experts and ordinary consumers and so putatively count as markers of the *artistic salience* of the formal, expressive, and semantic features of artworks within a community. The thought is that artworks are a natural kind of sorts defined by a common capacity to trigger aesthetic responses in consumers, however that might work. This kind of view is what philosophers call *an aesthetic theory of art*. Aesthetic theories of art collapse the concept of artistic salience into the concept of *aesthetic salience*, or reduce the former to the latter if you prefer. Empirical aesthetics associates the aesthetic salience of artworks with a capacity to produce a robust, objectively measurable, unified class of physiological responses in consumers. This, in turn, putatively warrants replacing the judgments and intuitions of experts about the nature of art with objective measures of the physiological responses of groups of ordinary art consumers. Perhaps no one really holds the kind of strong non-cognitive aesthetic thesis expressed here. Nonetheless, the objective measures of arousal used to evaluate judgments of artistic salience in empirical aesthetics are aesthetic measures (see Berlyne, 1974; Chatterjee, 2012; Silvia, 2012). So, on the surface of things, it looks as if empirical aesthetics relies on an aesthetic theory of art. The trouble is that a broad range of contemporary artworks are explicitly non-aesthetic, and, in the case of aesthetic works of art, there is no guarantee that the aesthetic responses of average members of an artistic community will track the correct use of the conventions governing artistic judgments and behaviors within that community.

1.1 Aesthetics and the philosophy of art

Discussions of the role of aesthetics in our engagement with artworks and the nature of artistic salience point towards a need for a *cognitive neuroscience of art*. The terms *aesthetics* and *philosophy of art* are each used in ordinary contexts to refer to philosophical studies of art. However, these terms are not synonymous and the equivocation in this casual, non-reflective use of language is significant. The category of aesthetics is broader than the category of art. It includes snowy mountain vistas, stormy
coastal landscapes, cozy capes and cottages with tidy lawns, industrial machinery, automotive design, and even the coffee pot on your kitchen counter. It also includes a range of artworks. The category art is broader than the extent of this limited range of artifacts. It includes formalist aesthetic Abstract Expressionist paintings and aesthetically striking Flemish landscapes from the 17th Century. But it also includes a full range of non-aesthetic conceptual works like Robert Morris' ephemeral installations, (e.g., Threadwaste, 1968) and Steam, 1968/9), Robert Rauschenberg's "combines" (e.g., Monogram, 1955-59), Jasper John's paintings of numbers and letters (e.g., 0 through 9, 1961), John Baldessari's paintings composed of text and definitions (e.g., Tips for Artists Who Want to Sell, 1966-68), and Sol LeWitt's Location Drawings (e.g., Wall Drawing 305, 1977). The concepts art and aesthetics therefore track different dimensions of the range of artifacts that we categorize as artworks and philosophy of art represents a broader, more inclusive definition of art. This is critical. It is certainly the case that some artworks are constructed for the purpose of displaying their aesthetic properties. However, many more, and likely the majority, of contemporary artworks are not like this. They are rather conceptual works. They express a concept or idea, challenge our concept of art, or perhaps reveal something to us about underlying biases with our communities (e.g., the works of John Baldessari, Joseph Kossuth, Robert Barry, Robert Morris, Wesley Meuris, The Judson Dance Theater Workshop, Barbara Krueger, Jenni Holtzer, or Adrian Piper). The trouble is that these works are not intentionally designed to trigger aesthetic experiences in consumers. In fact many of them are expressly anti-aesthetic, designed to draw attention to themselves as the ordinary, non-aesthetic objects, events, or actions that they are. Aesthetic theories of art would exclude these works from the category "art." Therefore, if empirical aesthetics depends upon a tacit aesthetic theory of art, it represents a narrow view of art that is inconsistent with contemporary practice.

A similar difficulty arises for aesthetic artworks. Recall that the range of aesthetic objects, scenes, and etc. is broader than the range of artifacts we classify as artworks. What then distinguishes artistic from non-artistic aesthetic objects? Consider Danto's discussion of Warhol's Brillo Boxes as a limit case (see Danto, 2000a). The original Brillo box is an aesthetically rich marketing device designed by a second
generation Abstract Expressionist painter named Steve Harvey. Harvey created the canonical Brillo box while moonlighting in commercial design to make a living. We can perceptually discriminate Warhol's facsimile's from the original by their relative size (Warhol's are larger) and material (Warhol's are made of plywood). But these visible differences don't alone suffice to disentangle which is the artwork. We could, to lapse into philosophical fantasy, imagine a Martian, or even a pigeon with similar perceptual capacities to humans (see Danto, 2003), mistaking Harvey's for a Warhol -- for that matter I suppose that we could imagine a naïve consumer with an adequate understanding of Pop Art but limited knowledge of Warhol's works within this category making the same mistake. Further, given that Warhol's Brillo Boxes reproduce the original visual design fairly accurately, they share the same range (or at least exhibit similar) dynamic, colorful aesthetic properties as Harvey's. The two types of artifacts aren't just visibly indiscernible in this regard, they are arguably also aesthetically indiscernible. What differentiates them as exemplars of different categories of artifacts, one an artwork and the other a commercial design object, is therefore neither their visible not their aesthetic properties. Danto argued that what enables us to categorically disambiguate these two sets of artifacts is the semantic salience of their visual design, or how these formal aesthetic properties were used to convey the content of the works in either case. Harvey's was a red, white, and blue visual celebration of the joy of Brillo intentionally designed to entice us to purchase a cleaning product. Warhol's was a Pop Art conceptual challenge to the dominant, formalist Abstract Expressionist aesthetic of the time. Or, the artistic salience of the formal aesthetic properties of Warhol's Brillo Boxes lies in what their hard edged formal design revealed about the then dominant conventions governing artistic practices within the New York avant garde art community. The aesthetic features of Harvey's Brillo box, in contrast, simply don't have artistic salience.

Warhol's work is conceptual, an expressly anti-aesthetic example of Pop Art. Someone might argue, consequently, that this is a bad test case. However, the model Danto proposes applies equally to works of pure abstraction as well. Consider, for instance, Jackson Pollock's One: Number 31, 1950. Ordinary consumers are struck by the energy of Pollock's monumental canvases, by their chaotic construction, perhaps exemplified in their failure to perceptually resolve into stable geometric patterns.
But the standard interpretation of these works locates their artistic salience in the trace of the gesture of Pollack's brush stroke, in the expressive activity of the artist and the way this aspect of the paintings self-consciously reflects Clement Greenberg's reductive formalist aesthetics (Greenberg, 1960) and Harold Rosenberg's concept of *action painting* (Rosenberg, 1952). The artistic salience of these works therefore lies in the fact that they are exemplars of a mid-Twentieth Century theoretical position about the nature of art exemplified by the *Abstract Expressionism* school of painting in New York City and the lives of the artists who defined it.

The discussion of *Brillo Boxes* and *One: Number 31, 1950* reveals something important about the nature of artworks: they are communicative devices. We typically engage artworks as intentional objects with cognitive content, as the product of a range of formal and compositional choices made by the artist, choices that carry information about their content. We are interested in how they were constructed, or perhaps better, in why they were constructed the way they were -- even when they are canonically aesthetic artworks. It can, as a result, be argued that the artistic salience of the formal-compositional features of a work lies in their semantic salience, in how they have been used to express the meaning or content of the work. Of course, artists often use aesthetic devices as part of the array of formal-compositional strategies they use to convey the content of their work. But, the artistic salience of these features lies in their semantic salience, in what they contribute to the content or meaning of the work. This is true of the broad range of artifacts that count as artworks, from anti-aesthetic conceptual works to those limit cases whose content simply involves the expression of some set of artistic conventions for constructing aesthetic artworks. For instance, the artistic salience of landscape paintings lie, at least in part, in what the style of painterly expression used reveals about the aesthetic conventions of the artist's time (e.g., the way the skewed scale of Hudson River School paintings reveal mid-Nineteenth Century American beliefs about the relationship between man and nature and the current politics of American Expansionism).

A standard philosophical objection to aesthetic theories of art can be used to illustrate this point. The trouble for aesthetic theories of art lies in the fact that they define artworks relative to a capacity for
producing aesthetic experience. This entails that where an artwork doesn't succeed in triggering an aesthetic experience in consumers, it isn't an artwork. If sound this would mean that that there could be no failed artworks. This, in turn, would mean that generations of folks who have toiled away in studios, struggling in vain with the creative process, were in fact not engaged in art making at all, but something else. Of course philosophical arguments with absurd conclusions like this one are rarely intended to be taken literally. The point of a philosophical *reductio ad absurdum* is more often than not to suggest that there is something wrong with some aspect of the reasoning in an argument that deserves attention and modification. "Art" is a success term. It's ordinary usage carries some degree of normative, or evaluative force. Some artifacts manage to "achieve" arthood, and we value them to the degree that that they do. The question is, what are the criteria that govern our judgments in this regard. One thing that can't do the work is a capacity to successfully trigger an aesthetic experience. Failed aesthetic artworks are artworks. They achieve the normal status of an artwork even though they fail to meet this desiderata. How do we recognize failed aesthetic artworks? We can see that they were designed with an intention to function as aesthetic objects, that they are fit to some set of formal-compositional conventions for producing works in that category of art, say aesthetically nuanced (as opposed to hyperrealist) landscape paintings, even if the formal-compositional relationships that they embody have failed to make the grade. What matters are the semantic properties that define the artistic content of the work and the fit of that content to the conventions governing categorical judgments about that class of artworks, i.e. that it is an artifact designed to be engaged with as an aesthetic object, and not anything about the actual aesthetic responses the work induces in consumers.

There is nothing particularly special about art in this regard. Artifacts achieve sawhood (or any sort of toolishness) when they meet the categorical criteria for being a saw (or any sort of tool). Likewise we value them for their success in achieving this status. What makes something a saw is that it has a blade suited for tearing some material. A saw blade may fail to realize this function. Toothless friction saws designed for precision metal cuts are particularly bad at cutting wood. Cross-cut saws are ill suited for making rip cuts. But we still categorize them as saws -- a novel type of cutting tool, say a length of rope
coated in an experimental abrasive designed to cut plastics would be a saw nonetheless, even if the abrasive wore away too quickly to be effective, by virtue of its design as a blade for tearing material. Likewise, a Claus Oldenburg soft saw sculpture is a saw of sorts even though it would fail to realize this function in any context.

1.2 Art, aesthetics, and appreciation

We might dub the issue discussed in the previous section the problem of artistic salience. The problem of artistic salience points towards a second philosophical challenge to empirically minded theories of art. Some philosophers have argued that questions about artistic appreciation, evaluative question about whether a work is apt to the conventions governing artistic practice, and if so whether it has been done well or poorly, are the critical questions that ground an understanding of art. This seems like too strong a position. Philosophers have always been interested in more general ontological and epistemological questions about the nature of artworks and our interactions with them. Aristotle, for instance, wrote on the narrative structure of tragedy and the nature of our psychological engagement with characters in The Poetics. There are similarly a range of interesting questions about the nature of installation art that are independent of questions about the value we assign particular installations (Irvin, 2013). What are the identity conditions governing an installation. Is an installation, like a musical score, a repeatable recipe? How should the installation crew implement the relationship between a particular installation and the particular space of the gallery? What if the artist doesn't leave explicit instructions, is vague about them, or simply isn't around to query for advice? Should the work stay as true as possible to relevantly similar prior installations? If so, which ones? If not what sort of constraints govern the creative judgments of the curator? These aren't evaluative questions. They are questions about the way artworks work, questions about the conventions governing artistic production, understanding, and related psychological processes. Nonetheless, evaluative questions are among the questions a philosophical or psychological theory of art should answer. This is a place where philosophy of art and empirical aesthetics agree -- aesthetic appraisals play a critical role in empirical aesthetics. The objective physiological measures of arousal that
ground research in the latter field are, in keeping with an aesthetic theory of art, explicit measures of artistic success, evaluative measures of aesthetic interest and appraisal.

Evaluative questions germane to artistic appreciation are *normative* questions, questions about the fit between an artwork and standard conventions governing artistic practice, or whether it was done well or poorly. The strong philosophical claim in this context is that causal explanations of behavior -- explanations of behavior in terms of the causal-psychological processes that link stimuli to behavior -- are irrelevant to normative questions. They apply equally whether or not that behavior is an appropriate response to the target situation or not (e.g. whether the arousal measures indicative of artistic interest and aesthetic responsiveness are appropriate to the work or are the product of contextually appropriate conventions and cognitive processes or not). For instance, we can imagine a consumer having a physiologically appropriate aesthetic response to a Jackson Pollack because it reminds them of the innocent exploratory creativity children's finger painting (perhaps their own child's finger painting).

Explanations of aesthetic appraisals in experimental psychology are causal explanations. Therefore, it is argued, they are not sensitive to the appreciative dimension of artistic behavior (Wittgenstein, 1966). Imagine, for instance, a study in which participants are asked to match musical compositions to a range of adjectives that expert music scholars and critics have identified as appropriate to those works (Dickie, 1962). Appealing to agreement among subjects here would be like appealing to agreement among young children about the meaning of nonsense words, or appealing to agreement among toddlers about the syntactic well-formedness of sentences. What really matters in these contexts is something about the conventions governing linguistic practices and how apt the stimuli are to those conventions. This information would already be in place at the start of this kind of study rendering its results otiose at best. Likewise with preference judgments. The preferences of uneducated novices are not the right kind of evidence to use to evaluate a work. The right place to look is the reflective judgments of experts about the fit between the work and the conventions governing those practices, not the averaged behavioral response of large samples of novice and expert consumers.

Appreciative judgments about artworks are grounded in comparisons between those artifacts and
the conventions governing artistic practices for a particular category of art. The success of these judgments rests upon an understanding of that category of art, including knowledge of the normative conventions that determine whether an artwork of that type was done well or poorly (Levinson, 1992; Rollins 2004). Philosophical questions germane to artistic appreciation are questions about the content of these conventions and why they have normative force, or how they constrain our evaluative judgments. These are not questions that can be answered by averaging the behavioral responses of ordinary untrained participants loosely associated with an artistic community. Consider the analogous case of morality. Ethical questions about moral behavior, questions about how one ought to behave in order to comport oneself as a genuinely moral person, can not be answered by observing the ordinary day-to-day behavior of members of a community or polling their preferences. These kinds of studies may reveal how groups of people do behave in a community, but the question about how one ought behave is a different question (see Moore, 1903). Objective measures of aesthetic interest and appraisal might similarly reveal what kinds of artistic judgments groups of people do make. However, these are objective measures of a participant's subjective response to an artwork. They do not differentiate between judgments that reflect an appropriate understanding of particular artworks and those that do not. The argument, therefore, is that it is a methodological error, a category mistake, to use behavioral methods and results to address questions about the structure and function of artistic conventions.

1.3 A model for a cognitive neuroscience of art

I have suggested elsewhere that philosophical skepticism about the utility of empirical research to any understanding of art is misplaced (Carroll, Moore, and Seeley, 2012; Seeley, 2011a; 2011b; 2008; Seeley and Kozbelt, 2007). Cognitive science, in the broadest sense of the term, refers to any of a range of interdisciplinary approaches to understanding how an organism acquires, represents, manipulates, and uses information in the production of behavior. Research in neuroscience of art and empirical aesthetics falls squarely within cognitive science under this interpretation. Artworks are stimuli intentionally
designed to trigger a range of affective, perceptual, and semantic responses in consumers constitutive of their artistically salient expressive, formal-compositional, aesthetic, and cognitive content. Empirical aesthetics and neuroscience of art are methodological tools that are regularly used to model the way consumers acquire, represent, manipulate, and use information carried in the formal-compositional structure of painting, films, dances, musical performances, novels and poems (to name a few) in the processes underwriting canonical cognitive, expressive, and aesthetic artistic responses to artworks. These data can, in principle, be used to adjudicate between competing philosophical theories about the nature of art and our engagement with artworks and answer thorny questions the problems of artistic salience and artistic appreciation raise for empirical aesthetics more generally.

A sketch for a general model for a rapprochement between philosophy of art and empirical aesthetics emerges from some of the platitudes grounding research methods in neuroaesthetics and a fairly standard story about artists' productive practices. The local environment is replete with information about the structure, function, and dynamics of objects and events, only a small fraction of which is germane to our current goals and activities at any given time. Perceptual systems, on the other hand, are limited capacity cognitive systems. Selectivity is therefore a critical component of perception -- perceptual systems must enable an agent to select task salient perceptual features and discard the vast range of distracting information that clutters his or her perceptual world. Recent research suggests that one means to solve this problem is to develop visual routines and other types of perceptual strategies for directing attention to sets of features diagnostic for the identities, affordances, and locations of task salient objects in the local environment (see Hayhoe and Ballard, 2005; Rollins, 2004; Schyns, 1998). Perceptual systems can, in this regard, be interpreted as evolved mechanisms for identifying and selecting task salient information from a dynamic flux of sensory inputs and discarding distractors in the local surround (Zeki, 1999; see also Pessoa, Kastner, & Ungerleider, 2002; Stein, Stanford, Wallace, Vaughan, and Jiang, 2006). Artists have developed systematic methods for selecting sets of diagnostic features from ordinary perceptual experience, rendering them in a medium, and evaluating their salience as productive strategies, e.g. color studies for paintings or story boards and dailies for film. The resulting formal-
compositional devices work as communicative strategies because they are directed at the range of perceptual strategies cognitive systems employ in ordinary contexts to smoothly cope with a dynamic environment replete with sensory information on the fly. Therefore we can hypothesize that we ought to find a close coupling between artists' productive strategies and the operations of perceptual systems.

A significant amount of research in neuroaesthetics has been devoted to uncovering, tracking, and evaluating the range of ways artists' formal-productive strategies map to the operations of perceptual systems (e.g., Richard Latto's discussion of the relationship between irradiation and lateral inhibition in the retina, Beatrice Calvo-Merino and Corrine Jola's explorations of the role played by premotor areas in the action-observation-network in audience engagement with dance, and Nicole Speer's research on the role played by motor simulation in narrative understanding in text and film) (Calvo-Merino et al, 2008; Jola et al, 2012; Latto, 1995; Speer, 2009). However, correlations between the neurophysiological operations of perceptual systems and artists' formal-compositional strategies do not alone suffice to explain the artistic salience of the formal, expressive, or semantic features of an artwork. Irradiation works to enhance figure ground segregation and the perception of depth in Monet's paintings because lateral inhibition produces half-shadows everywhere in our perceptual environment that play an analogous role in depth perception and object recognition (Latto, 1995). Motor simulation, likewise, works as a mechanism to support the role of kinetic transfer in dance communication and narrative understanding in film because it plays an analogous role in action recognition and the perception of biological motion in ordinary contexts (see Carroll & Seeley, 2013; Seeley & Carroll, 2013). Similarly, motor simulation plays a role in narrative understanding in fiction because (or to the degree that) it plays a critical role in the semantic comprehension of action sentences (Fischer and Zwann, 2008; Speer, 2009). What one needs is a further argument connecting the way artworks work generally as cognitive stimuli to the way consumers recognize and understand the artistic salience of their formal, expressive, and semantic content. The notion of an aesthetic appraisal is often called on to do the heavy lifting here in empirical aesthetics. However, as discussed above, the preponderance of non-art aesthetic objects, events, and experiences raise a question about whether aesthetic measures are appropriate measures of artistic salience.
A solution to this problem emerges from the observation that there is no ideal formal-compositional strategy for artistic production in a medium. Rather artists must choose how to render their subject matter from a wide range of alternatives. Categories of art in which productive practices are directed towards shared subject matter are the easiest place to see this, e.g. realistic landscape painting and Hollywood movies, but any exploration of formal stylistic differences would ultimately do. The stylistic variance in what has passed as realistic depiction in this context demonstrates that any of a range of formal-compositional strategies suffice to render these environmental features realistically (see Gombrich, 1960; Lopes, 1995). We might tell a similar story about the range of variants of common stories in movie and theater production, e.g. Pierre Choderlos de Lacos 18th Century novel Dangerous Liaisons, Christopher Hampton's theatrical version of the same story, or Stephen Frear, Jin-ho Hur, and Milos Foreman's film versions. What are the constraints governing the development of those formal-compositional strategies that define the artistic styles of different schools, eras, and artists along with the surface structure of different artworks -- the intended content of the work, the expressive, formal aesthetic, and cognitive effects an artist intends the work to have on consumers. What are the constraints on the artistic content of a work? The conventions governing productive and appreciative practices for a category of art. Therefore, the formal-compositional structure of an artwork naturally carries information about the conventions governing artistic practices, information that cognitivist models within the philosophy art argue determines the artistic salience of the expressive, formal aesthetic, and semantic features of a particular work. This entails that artworks are attentional engines, or cognitive stimuli intentionally designed to direct attention those aspects of their formal-compositional character that carry information about their content, including information about the category of art a work belongs to (Carroll and Seeley, 2013; Levinson, 1992). Categorizing an artwork as belonging to a particular school, era, or artistic style provides a recipe for how to approach it -- instructs a consumer as to which sets of artistic conventions are appropriate to his or her cognitive and appreciative engagement with the work (e.g., as a satire or a dramatic novel, an Impressionist or a Hudson River School painting, or a Modernist or a Postmodernist dance).
A core set of critical questions loosely define this research model for philosophy and neuroscience of art (Seeley 2013). What are the range of formal-compositional strategies that define productive practices within a category of art? How do these strategies function to carry and communicate diagnostic information about the content of these artworks? What accounts for the artistic salience of the expressive, formal aesthetic, and semantic features constitutive of the content of these works? A solution to the problem of artistic appreciation follows naturally from the answer to the last question. The artistic salience of the expressive, formal-aesthetic, and semantic features of an artwork lies in the way they reflect the range of artistic conventions that govern artistic practices for the category of art to which it belongs. Danto's discussion of Warhol's *Brillo Boxes* is a canonical illustration of this position. The artistic salience of Warhol's *Brillo Boxes* lies in the fact that they violate, and so cause us to reflect on, the conventions governing productive practices within, and appreciative judgments about, Abstract Expressionist artworks. Appreciative judgments about Warhol's work involve judgments about the validity of this project in the context of conventions governing avant-garde art, which are conventions governing the reflective use of artistic practices to evaluate the validity of artistic conventions. Warhol chose to approach this problem using a visually pleasing design. But, ultimately his strategy was to highlight the identity of the work as what it was, an exemplar of an ordinary household object whose salience was tied to its everyday use, not a metaphysically charged, abstract, high-art aesthetic object. Other examples of this strategy are Jasper Johns' *Savarin Coffee Can with Paint Brushes* (1960) and *Painted Bronze (Ballantine Ale)* (1960) and Steve Paxton's untitled dance in which he sat on a bench and ate a sandwich (see Banes, 1983). We might argue that the visually striking features of the Brillo design are, in this regard, the least successful aspect of Warhol's work! Alternatively we might argue that he employed a visually striking design to draw attention to its irrelevance to the salience of the artifact as an artwork.

2. Hierarchical Models and Attentional Engines

We need a cognitive neuroscience of art because we need a model for how we use semantic content to
disambiguate artworks from similarly structured non-art artifacts, e.g. works of commercial design. This is transparently obvious in the case of non-aesthetic conceptual art. But, it is also true of abstract aesthetic works at the opposite end of the spectrum. The artistically salient content of a work emerges from the way it's formal-compositional structure reflects the conventions governing artistic practice within the appropriate category of art. The canonical example used to illustrate this fact within philosophy of art comes from Kendall Walton (1974). Imagine two perceptually similar categories of art, Picasso's *Guernica* and a hypothetical category of *bas reliefs* that share its formal-compositional structure called *Guernicas*. The artistically salient dynamic features of *Guernica* are a product of its fractured, dynamic, black and white cubist composition (which Picasso used to express something about the horror of war for the communities within which they are fought). *Guernicas* are, in contrast, colorless. Their artistically salient dynamic features emerge from the way figure ground relations give way to a depiction of the chaos of war in their three dimensional structure. *Guernica* is flat. Therefore, relative to the conventions governing bas reliefs it is static rather than dynamic. The same can be said of *Guernicas*. They are colorless. Therefore, relative to the conventions governing artistic production in painting they lack an articulated structure sufficient to support the dynamics of *Guernica*. We need a cognitive neuroscience of art because we need a theory that is sensitive to the way these kinds of conventions govern artistic practice within a category of art, constrain the contents of artworks, and influence our engagement with them. We need a theory that can explain how artworks work as attentional engines to focus attention on formal-compositional cues diagnostic for the appropriate category of art, e.g. *Guernica* and not *Guernicas*, *Pop Art* and not *commercial design*, or *Abstract Expressionism* and not *finger painting*.³

Biased competition theories of selective attention suggest a model for how artworks function as attentional engines. Recall the problem of selectivity. How do perceptual systems solve this problem, how does an organism flexibly orient its attention to diagnostic information on the fly in the kinds of novel contexts it confronts in everyday behavior? The simplest solution is that these systems are fine tuned to *perceptual salience*. Some environmental features are simply brighter, more vividly colorful, or move differently than their neighbors. We easily orient our attention to them because they stand out in a crowd.
The trouble is that the vast majority of behaviorally salient features in the environment simply are not perceptually salient in this way -- certainly none of the keys on my keyboard, the papers on my desk, or the baking products in my cupboard sing out their relative salience to my current needs like a klaxon. Therefore perceptual systems require an independent mechanism for assigning salience to diagnostic features -- for tagging the salience of objects and their features with relevance to me rather than their relative difference from their surround. Biased competition models suggest that we employ top-down fronto-parietal and cortico-fugal attentional networks to solve this problem in everyday perceptual contexts. These networks bias perception by priming the firing rates of populations of neurons in sensory systems to the expectation of diagnostic features at particular locations. A fast first sweep through the perceptual system is sufficient to match minimal diagnostic features to semantic knowledge of the structure, function, and affordances of object and event types stored in declarative memory (Pessoa & Adolphs, 2010). This information, in turn, can be used to scaffold perception -- to direct attention and bias perceptual processing to task salient features, object or object parts at expected locations within the visual field.

We can tell a similar story about artworks. Some of the features of an artwork are perceptually salient, and so naturally draw our attention. For instance Winslow Homer rendered the trigger hand of the Union sniper brighter than any other feature in his wood engraving (for Harper's Weekly) The Army of the Potomac - A Sharp-Shooter on Picket Duty (1862). The morality of the practice of using sharpshooters during the Civil War was questionable. These soldiers never faced the enemy they engaged and so it was thought that the use of sharpshooters dehumanized the practice of war. We can't see anything of the face of the sharpshooter in the print beyond a shadowed eye trained on his target down the barrel of his rifle. Our full attention is drawn to his hand on the trigger (which hides his face). The sharpshooter is, thereby, depicted as a dehumanized mechanism of destruction. Likewise, the raw stroke of yellow oil crayon John Singer Sargent used to depict the fishing rod in Val d'Aosta: A Man Fishing (1907) and the red flowers of the poppies depicted in Monet's Wild Poppies Near Argenteuil (1873) are perceptually salient features. However, the artistic salience of these works does not lie in the either the coarse structured yellowishness...
of the fishing rod or the bright rough redness of the poppies. Rather these features function to draw our
attention to more general structural features of the works, the dynamic interplay of light on the surface of
the pond beside the fishermen in the Sargent and the generically greenish-brown field of grass waving in
the wind that dominates the lower half of the Monet. In fact the whole composition of the Sargent
painting -- the orientation and posture of the prone figures lying beside the pond, the gaze of the
fisherman, and the orientation of the fishing rod -- drives the consumer's attention to the coarse expressive
brushwork constitutive of the dynamic depiction of the water. These structural features are diagnostic
cues to the identities of the two works as exemplars of an Impressionist style of painting. These cues in
turn let us know how to engage the work, in this case to pay attention to the way the brushstrokes are used
to depict the dynamic interplay of light, form, and movement in the visual field.

What is the mechanism in all of this? Biased competition models of selective attention identify a
broad network of fronto-parietal and cortico-fugal circuits that facilitate the crossmodal integration of
sensorimotor and unimodal sensory information in perception. This network includes projections from
prefrontal areas involved in spatial working memory, e.g. dorsolateral prefrontal cortex (dLPFC) to the
frontal eye fields (FEF) which are associated with endogenous shifts of attention, and projections from
both FEF and dLPFC to superior colliculus (SC), pulvinar, and the lateral geniculate nucleus (LGN) in the
thalamus (Kastner, 2004). These subcortical areas are hypothesized to participate in processes critical for
neural synchronization (LGN and pulvinar), directing eye movements (SC), and the crossmodal
integration of visual, auditory, and somatosensory information (SC). LGN is the primary relay station
between the retina and the visual cortex. Therefore, attentional priming reaches back to the earliest stages
of visual processing (Beck & Kastner, 2009; Stein et al, 2004; Pessoa & Adolphs, 2010). These feedback
projections from prefrontal and parietal areas that are associated with object recognition, spatial working
memory, and attention to topologically encoded cortical and subcortical perceptual areas facilitate the
integration of sensory information into coherent representations of task salient aspects of the distal
environment and account for the influence of categorization processing in perception. Feed forward
projections from perceptual processing areas to these same prefrontal and parietal areas, in turn, function
to reinforce perceptual representations and instantiate working memory in a distributed multi-modal network.

The trick is that our engagement with paintings depends on a capacity to reconstruct their depictive and representational content from diagnostic information encoded in the abstract sets of marks constitutive of their two dimensional formal-compositional structure -- to recognize objects and scenes in a painting and comprehend what it would have meant for an artist to render them in the particular way they have been painted (e.g. the coarse brushstrokes used to depict the pattern of reflected light that gives the water its dynamic appearance in Sargent's *Val d'Aosta: A Man Fishing*). This requires matching abstract diagnostic cues both to category knowledge in object recognition and knowledge of the stylistic conventions regulating artistic production in a category of art -- or, recognizing an artwork both as a visual representation of what it depicts and as a member of a category of art. The suggestion is that cortico-fugal attentional circuits (in conjunction with a range of further recurrent connections, e.g. projections from areas TE and TEO to visual areas V4 and MT associated with the perception of color, abstract patterns, motion and depth) play a critical role in this capacity.

Hierarchical theories within empirical aesthetics can be used to articulate the range of processes that might influence cognitive and appreciative judgments about the artistic salience of the content of a work. These theories model artistic production and our engagement with artworks as a multi-stage process that includes sensory processes, perceptual analysis, explicit classification judgments, decision procedures, and affective responses that underwrite our capacity to recognize the formal-compositional and depictive content of a work, understand its semantic-representational content, and appraise its artistic salience (Chatterjee, 2003, 2012; Kozbelt & Seeley, 2007; Leder et al, 2004; Nadal et al 2008). For instance, Leder and his colleagues differentiate between two distinct kinds of aesthetic responses to artworks, *aesthetic judgments* and *aesthetic emotions*. Aesthetic judgments are a product of what they call *cognitive mastering*, or the explicit classification and subsequent interpretation of the content of a work relative to expert knowledge of stylistic conventions that define artistic schools, movement, eras and the works of individual artists, e.g. Rothko and Motherwell's Abstract Expressionists works differences
between Sargent's portraits and his Impressionistic landscapes. Explicit classification is the key in these processes. It specifies the set of artistic conventions that constrain our cognitive engagement with a work. The output of perceptual analyses can then be meshed with salient expert art critical knowledge to yield an interpretation of the meaning of the work in cognitive mastering. A reduction of ambiguity, or uncertainty cognitive mastering (or lack thereof), then yields an aesthetic evaluation of the success of a work. Success at the explicit classification, cognitive, and evaluative stages are experienced as changes in affective state that yield a state of pleasure or satisfaction Leder defines as an aesthetic emotion.

Leder argues that content responsiveness will lie of a continuum from naïve to expert consumers (see also Cela-Conde, Marty, Munar, Nadal, & Bruges, 2002). Naïve consumers who lack expert art critical or art historical knowledge will respond to the depictive content of a work, evaluating them relative to their success as representational works or abstract constructions on this model. Expert consumers, on the other hand will exhibit sensitivity to the artistically salient formal, expressive, and semantic features of the work and the way these are used to generate its art critically salient content. This is precisely what our model would predict. Loosely explicit classification and cognitive mastering are processes that involve the implementation of judgments about the relative diagnosticity of features of the artwork for its content. these processes in turn, bias processing in fronto-parietal and cortico-fugal attentional circuits, driving our perceptual, affective, and cognitive interaction with the work (see also Chatterjee 2003, 2012) discussions of our engagement with artworks.

3. Artistic salience, artistic appreciation, affective and embodied responses to artworks

We can now return to the questions about artistic salience and appreciation that have driven philosophical skepticism about empirical aesthetics. The first thing to note is that hierarchical theories have resources ready to hand to tackle these issues. These theories model our engagement with artworks as a multi-stage process that employs expert art critical knowledge to integrate perceptual, cognitive, and affective aspects of our engagement with artworks. Other theories model our engagement with artworks as a cognitive process, e.g. prototype preference (Martindale and Moore, 1988) and processing fluency theories (Reber,
The virtue of hierarchical theories over these other approaches is that nested sets of hierarchically related recurrent circuits enable them to explicitly model the way the formal-compositional features of a work achieve artistic salience -- the way they carry the information about salient artistic conventions constitutive of the formal aesthetic, expressive, and semantic content of a work. Nonetheless, and this is the real rub, all of these theories model the influence of cognitive variables in our engagement with artworks as a part a process of aesthetic appraisal -- they, like aesthetic theories of art more generally, treat artworks as *aesthetic engines*. For instance, Chatterjee (2003,2012) argues that *aesthetically* salient perceptual, expressive, and semantic features of artworks drive top-down processing and attentional circuits in our engagement with artworks. This, as discussed above, narrows the potential explanatory scope of these models and opens the door to the problems of artistic salience and appreciation. However, if we drop the aesthetic desiderata and treat artworks as attention engines instead, these problems fall away as spurious worries.

The problem of artistic salience is a question of what differentiates artworks from non-art artifacts. This is a philosophically important question. A minimum desiderata for a theory of art is that it explain how we sort the range of artifacts that fall under that category, how we discriminate artworks from non-artworks and sort them into salient categories of art. Theories in neuroaesthetics, and empirical aesthetics more generally, treat artworks as cognitive stimuli like any other and model our engagement with art as a product of everyday cognitive processes. The challenge for any model like this that treats artworks as ordinary cognitive stimuli is to explain then how we manage to sort the artworks from the other everyday perceptual stimuli. One common strategy in neuroaesthetics is to simply dodge this question and study the way a set of standard, successful formal aesthetic strategies work, e.g. discussion of irradiation and *McKay effects* in Monet's paintings (Latto, 1995; Livingstone, 2002). In these contexts researchers define artistic salience relative to aesthetic salience, take the aesthetic salience of a well accepted formal strategy as a given in a particular context, and then explain associated aesthetic effects as products of our perceptual engagement with the work. However, as discussed above, the artistic salience of the aesthetic features of a work of art is not a perceptual given, but rather depends on the category of
art to which we assign the work, to the way we differentiate it from other non-art artifacts and sort it from its other artistic cousins.

The diagnostic recognition framework suggested by biased competition theories of selective attention models our engagement with artworks differently. In ordinary perceptual contexts minimal sets of diagnostic features suffice to enable an organism to perceptually recognize the identities and affordances of objects and events (Kosslyn & Koenig, 1995; Schyns, 1998). How might this work? A fast first sweep through unimodal perceptual systems might suffice to match sparse sets of coarse coded sensory stimuli to declarative knowledge of the structure and function of categories of objects or events stored in semantic or episodic memory. This knowledge can, in turn, be used to direct attention and prime the firing rates of populations of neurons across the full range of perceptual modalities to other environmental features diagnostic for the task at hand. These processes enhance the encoding of task salient perceptual information, inhibit the encoding of potential distracters, and thereby facilitate the smooth and efficient selection and perception of behaviorally salient information. Categorical knowledge of the structure and function of object and event types therefore serves as a strong determinative constraints that gates sensory processing, influencing not only what we see, but how we see it (Pessoa et al, 2002; Schyns, 1998).

Artworks can be sorted into unique categories of objects and events defined by media (sculpture vs. painting), productive practices (dance vs. performance), content (e.g. portraits vs. landscapes; formalist vs. conceptual), and formal-compositional stylistic features (e.g., Thomas Cole's *Hudson River School* paintings, Diego Rivera's *Social Realism*, Thomas Hart Benton's *Regionalism*, or Robert Bechtle's *Hyperrealism*). These categories of art, in turn, articulate sets artistic conventions that constrain our engagement with artworks. The identities of individual artworks may cut across categories of art -- or, categories of art may not sort into functionally discrete sets, but rather form an interlocking web, and integrated network that defines an artworld at a time. For instance, Bill T. Jones borrows elements of performance to articulate the narrative content of his choreographed works (e.g. his work about the life of Abraham Lincoln, *Fondly Do We Hope... Fervently Do We Pray*) and Judy Pfaff borrows from Cubism
and the painterly drawing style of Jackson Pollack in her sculptural installations (e.g., *NYC, BQE*, 1987 or *Low Muse*, 2010). Recognizing the productive role these shared categorical conventions play in the content of the work is critical to understanding them. A capacity to categorize the formal-compositional content of the work in this way guides attention, facilitates the recovery of diagnostic features, and, via the same attentional processes operative in ordinary perceptual categorization tasks, determines the artistic salience of what we perceive and experience in our engagement with artworks. Therefore, treating artworks as attentional engines and acknowledging the productive role of categorization processing in attention and perception transparently resolves the problem of artistic salience for empirical studies of the arts.

A solution to the problem of artistic appreciation falls our for free on this view of how we recover the artistic salience of the features and content of an artwork. Recall that aesthetic appraisals and other appreciative judgments about artworks are normative judgments about whether a work was done well or poorly, about whether it is apt to a context. Normative judgments of this sort involve comparisons between the formal, expressive, and semantic content of the work and conventions that govern salient artistic practices, e.g. conventions constitutive of artistic practices within the category of art to which a work belongs. Psychological explanations are causal explanations, or explanations of the causal-psychological interactions between a work and a consumer. The initial skeptical claim is that causal interactions between a consumer and an artwork are immune to the influences of these conventions. Psychological explanations of our interactions with artworks are therefore not sensitive to the right kinds of variables to stand as explanations of the success or failure of appreciative judgments -- they can't alone discriminate between normative judgments that are appropriate and those that are not. This simply isn't true of the model articulated here. From the perspective of a diagnostic recognition framework, our perceptual engagement with artworks is driven by attentional processes that are themselves governed by knowledge of the range of conventions that define different categories of art. These conventions are, as philosophers say, constitutive of way we experience the work. They are productive constraints that determine what we perceive in a work and how we perceive it. Where the work doesn't match to the range
of conventions appropriate to a category of art our perceptual expectations will be unsatisfied, our attentional strategies will misfire and remain unrealized. Our experience of the work will consequently be degenerate. It will fail to meet the cognitive goals that drive it and we will appraise it as ineffective.

The model I have proposed here, like hierarchical models generally, has the resources to resolve the problem of artistic appreciation because the salient artistic conventions are built into the perceptual processes that enable a consumer to recover the content of a work. Two caveats are in order here. First, a consumer's expert art critical knowledge can be incomplete, incorrect, or non-existent. In these kinds of cases we would expect that the appraisals they generate for artworks in those categories would not adequately track their artistic salience. But this is not a problem for the model. Fit to artistic conventions are the markers of artistic success on this account, not physiological responses and preference judgments. We, therefore, have the resources are ready to hand to evaluate inappropriate appraisals. Nonetheless, the skeptic may yet find this approach wanting. Everything turns on how one interprets the claim that questions about art are questions about the conventions governing appreciative judgments about artworks. The kinds of questions answered here are how questions about the role artistic conventions play in our engagement with artworks. They are not why questions about the reasons we find value these conventions in the first place. The short answer to the second question is, "...because they define the objects, events, and practices constitutive of our categories of art." But this answer won't be particularly satisfying to anyone who takes why questions seriously. My intuition is that we can appeal to dopaminergic reward systems to model the emergence of artistic value in our engagement with artworks as a how question. It is possible that an anthropological and evolutionary accounts of art could give us purchase on related why question, as might studies in cultural criticism at the humanities end of research on the arts. These are hard questions for future directions and another day though.

References:


Madison, WI: University of Wisconsin Press.


Pessoa, Louiz and Adolphs, Ralph (2010). Emotion processing and the amygdala: from a 'low road' to 'many


Endnotes:

2. See Seeley (in press) for a more detailed discussion of these points.
3. It is critical in this context to disentangle the notion of content from the related concept of interpretation. Interpretation is a tricky business. I take it that contextual factors influence the process of interpretation and constrain the validity of a particular interpretation on an occasion -- consider, for instance, differences in the interpretation of politically charged conceptual artworks as contemporary vehicles for social change as opposed to historical artifacts. The content of these kinds of works does not change across historical contexts. But an interpretation of the work as a live expression of current normative social issue may well be strikingly different than retrospective historical interpretations of the same work as an expression of a political movement embedded within the broader social structure of its time. The same might be said for the relationship between an artwork and an artistic movement -- surely the broad lens of history yields a different interpretation of action paintings like Jackson Pollack's than did the polemical lens of mid-century art criticism current to their time. This is not to say that interpretation is a subjective matter. Rather, interpretation must be flexibly sensitive to changes in social and epistemic context in a way that categorical judgments that constrain our understanding of the content of an artwork are not (see Danto, 2000b). Pollack's work is an expression aspects of Greenbergian
Modernism, facts about the flatness of its composition and the transparent roughness of his brushstrokes reflect Greenberg's attention to the formal aspects of medium and support, whatever it might have meant then or means to us now for someone to have painted that way.