Putting reward in art: A tentative prediction error account of visual art

Sander Van de Cruys
Laboratory of Experimental Psychology, Tiensestraat 102, 3000 Leuven, Belgium;
e-mail: sander.vandecruys@psy.kuleuven.be
Johan Wagemans
Laboratory of Experimental Psychology, Tiensestraat 102, 3000 Leuven, Belgium;
e-mail: johan.wagemans@psy.kuleuven.be

Abstract. The predictive coding model is increasingly and fruitfully used to explain a wide range of findings in perception. Here we discuss the potential of this model in explaining the mechanisms underlying aesthetic experiences. Traditionally art appreciation has been associated with concepts such as harmony, perceptual fluency, and the so-called good Gestalt. We observe that more often than not great artworks blatantly violate these characteristics. Using the concept of prediction error from the predictive coding approach, we attempt to resolve this contradiction. We argue that artists often destroy predictions that they have first carefully built up in their viewers, and thus highlight the importance of negative affect in aesthetic experience. However, the viewer often succeeds in recovering the predictable pattern, sometimes on a different level. The ensuing rewarding effect is derived from this transition from a state of uncertainty to a state of increased predictability. We illustrate our account with several example paintings and with a discussion of art movements and individual differences in preference. On a more fundamental level, our theorizing leads us to consider the affective implications of prediction confirmation and violation. We compare our proposal to other influential theories on aesthetics and explore its advantages and limitations.

Keywords: psychoaesthetics, Gestalt, art perception, predictive coding, reward, aesthetic emotion, perceptual organization.

1 Introduction

It was the Gestalt psychologist Koffka who stated that violations of the law of the good Gestalt "hurt our sense of beauty" (Koffka 1935, p 174). If we take "good Gestalt" (*Prägnanz*) to mean having regular, clear, symmetrical, and simple forms, as it is conventionally defined, this seems to exclude an abundance of artworks (ranging from traditional to modern) from ever being beautiful to human eyes. Why is it that artworks, prime instances of beauty, often contain precisely these violations that are supposed to "hurt our sense of beauty"?

We will suggest a way out of this conundrum based on a view on visual art that is firmly rooted in cognitive neuroscience. It will be less reductionistic than other proposals in that art will not be explained by piecemeal activations in separate visual areas on a specific level of visual processing but by taking into account the bidirectional, hierarchical nature of the visual system as it is implemented in the predictive coding approach. The former approach is doomed to fail for at least two reasons. First, an activation, however strong, of one or more of the visual processing areas is clearly not sufficient for an object or event to pass as a work of art (Hyman 2010). Melcher and Bacci (2008, p 357) rightly remark: "A particular artwork may indeed activate area V5/MT, but a passing bicycle would activate those motion-processing areas even more strongly and would not be considered a work of art." Second, we are indeed never confronted with the isolated raw visual input; rather, our visual system (and the brain as a whole) actively but effortlessly organizes it into surfaces, motions, three-dimensional objects, and concepts. Our proposal will center on these processes of perceptual organization (rather than the representational or symbolic meaning of their "end-products"). The rationale for this is found in Redies (2007, p 3), who calls for a broadly applicable theory

(from cave art to Kandinsky) "as universal as aesthetic judgment itself." He notes that since almost any (visual) stimulus can be used to compose aesthetically pleasing objects, aesthetic perception may rely on general aspects of information processing that are implemented in all visual channels and regions. Redies adds that aesthetic perception requires the processing of global features encompassing interactions between large amounts of receptive fields. By grounding our proposal in the formal, organizational characteristics of visual information processing as hypothesized by the predictive coding framework, we will attempt to meet these requirements.

Perceptual processes alone would hardly give us aesthetic experiences, though. Together with anthropologist Dissanayake, we claim that "much is overlooked when aesthetic cognition is conceptualized simply as 'sensory' or 'perceptual'..." (Dissanayake 2009, p 163). A one-sided emphasis on perception neglects the emotional aspects that are the motivating drives for creating and enjoying art in the first place. Artistic expressions attract and hold attention and stir and shape emotion. Consequently, any theory of art must explicitly elucidate the crucial interactions between perceptual and emotional processes. Luckily for our purposes, recent data suggest vision may be intrinsically affective (Barrett and Bar 2009), in the sense that the processing of emotional relevance and value is not an afterthought in the visual processing hierarchy but can actually drive object formation and recognition. Our challenge will be to specify the emotional consequences of the formal, organizational mechanisms of the predictive coding view, bracketing the emotional or symbolic content of the representations involved.

2 Predictive coding

The predictive coding approach of perception holds that the brain actively anticipates upcoming sensory input rather than passively registers it. On the basis of prior experience, the brain actively makes predictions about what visual input to expect in the current context of stimulation. At every level of the visual hierarchy predictions are generated and propagated (top-down) to lower levels, where they are checked against incoming (bottom-up) evidence. The idea is that these predictions suppress or explain away the activity in lower areas that agrees with them (de-Wit et al 2010), while what remains and is sent upward are the mismatches between these predictions and the current input, also called the prediction errors. This way the processing resources (attention) can be directed to that part of the stimuli that has not been sufficiently explained (predicted), and thus still has to be learnt. Through constantly fine-tuning predictions using the mismatches, the brain becomes tuned to statistical regularities of our natural visual environment. These predictions structure the perceptual input in patterns that allow for predictability both within and across visual displays. Thus, the classic concept of Gestalt, traditionally defined as a (experiential) whole that is different from the parts, can be recast in terms of predictive coding (Van de Cruys and Wagemans in press).

This framework is fruitfully adopted to explain several findings on visual processing (but also auditory perception [cf Winkler et al 2009; Kumar et al 2011]). It can account for extra-classical receptive field effects measured with single-cell recordings in the primary visual cortex (Rao and Ballard 1999) but also for fMRI patterns of activation across the visual hierarchy. For example, Murray et al (2004) observed that perceptual grouping is accompanied by an increase in activity in higher tier object-sensitive areas (LOC) and concomitant decreases of activity in lower visual areas (V1), as predicted by a predictive coding view. Furthermore, the well-known phenomenon of repetition suppression can be explained by a reduction of neural activity for predictable stimuli (Summerfield et al 2008). Similarly, Alink et al (2010) cleverly used an apparent motion path to generate strong

predictions about when and where a visual stimulus will be, demonstrating that stimuli evoke smaller responses in V1 when they have an onset time that matches these predictions. The more traditional theory of (neural activity in) perception as the piecemeal accumulation of evidence with visual neurons functioning primarily as feature detectors would have great difficulties in explaining this series of evidence. In contrast, in predictive coding, perception is an iterative matching process of top-down predictions checked against bottom-up evidence along the visual hierarchy. Consequently, each level in the visual cortical hierarchy has a twofold computational role: first, it provides predictions (the conditional probability of a stimulus) regarding expected inputs to the next lower level, and second, it encodes the mismatch between predictions and bottom-up evidence (the prediction error or "surprise"). Egner et al (2010) explicitly set out to compare performance of a predictive coding model with the traditional model. They found that when subjects strongly expected to see a face, fMRI activity in the face-selective fusiform area was indistinguishable when actually viewing a face versus a house, while maximally differentiated when the expectation of seeing a face was low. Using computational modeling, the authors conclude that this pattern of results can only be accounted for by the predictive coding model, which says that the total neural activity in category-selective areas represents the sum of activity related to prediction ("face expectation") and that related to prediction error ("face surprise").

At the heart of predictive coding is the concern for ease (efficiency) of processing. As described earlier, neural resources needed for processing predictable stimuli are minimized, as our system gradually becomes optimized to the statistics of our natural perceptual environment. In addition to that, the predictive coding strategy is parsimonious because different levels of the processing stream do not need to keep duplicates of information, maintained in higher regions (de-Wit et al 2010). In light of evolution, being able to encode and process sensory information in an efficient way is vital for the costly organ the brain is.

Being able to successfully predict is another obvious evolutionary advantage, as it allows animals not only to react after the fact to stimuli that change the internal milieu (homeostasis) but also to prepare (anticipate and compensate) for those that are very likely to ensue. Thus, homeostasis urges organisms that can walk around and manipulate their environment to take a predictive stance. Friston's (2010a) generalization of the predictive coding framework starts from homeostasis, or the realization that only a limited set of all the states an organism can be in is compatible with its continued existence. The long-term goal of reducing the time spent in "surprising" states translates into the short-term goal of reducing prediction errors.

This holds for single-celled organisms as well as for complex mammals, like humans, who evidently can rely on a greatly enhanced predictive capacity (Cerra and Bingham 1998). Based on statistical regularities in the environment, organisms form predictions on the where, when, and what of future resources, compensating for the inherent variability in the availability of these resources. Complete predictability implies that any disturbances of internal milieu are fully compensated for (by plasticity or action). Propagated prediction errors tell our system whether our current cognitive resources are up to the task of interpreting and coping with incoming stimuli or whether extra effort is needed. This effort can take the form of learning, when changing the predictions (our internal generative model of our environment), but it can also be behavioral effort, when actions are executed to change the things predicted or our sampling of it (Friston 2010a).

If we acknowledge that predictive coding is ultimately founded on homeostasis and the vital maintenance functions of the body, it should be no surprise that we can link this view to emotions. Indeed, to understand aesthetic emotion (appreciation), we need to make this link explicit.

3 Prediction and emotion

Emotions can be seen as motivational amplifiers (Huron 2006). They motivate organisms to pursue behaviors that are normally adaptive and to avoid behaviors that are normally maladaptive. When they form accurate predictions, organisms can efficiently react to upcoming events, thereby increasing the likelihood of future positive outcomes. Hence, it would be wise for evolution to reward cases in which predictions are confirmed in actual circumstances. Meanwhile, our failures in predicting situations may be characterized by negative emotion, because they signal that there is something wrong with the mental model we use to generate the predictions. It follows that prediction errors are always to some extent negative in affective valence. Adopting a better-safe-than-sorry strategy, nature tends to assume the worst because the cost of a false negative (type II error) is potentially much larger than that of a false alarm (type I error). In a conservative reflex, the first, quick reaction to prediction violation or error is negative (Huron 2006). This is also reflected by the so-called conflict theories of emotion (for a historical review see Mandler 2003), which claim that emotions arise from interruptions or discrepancies between expected and actual situations. Hebb (1949) and Mandler (2003) are the most well-known advocates of this view.

By studying the continuous intrinsic activity of the brain, Hebb realized that the brain was proactively involved in processing incoming stimuli, rather than just passively responding to them. According to him, thought consists of so-called phase sequences: sequential activations of neural structures (cell assemblies) that are built up as a result of previous experience and learning. Each assembly activation may be aroused by a preceding assembly, by a sensory event, or by both. Negative emotions arise, then, from the interference with (obstruction of) such an established phase sequence. Although the terms used may seem quite peculiar to modern ears, Hebb's view is consistent with the approach described earlier of implicit formation of predictions and their confirmation or obstruction.

Mandler (2003) elaborated on this theory, arguing that interruptions of ongoing response tendencies and conflicts between expectations and actual circumstances create arousal because they signal important changes in the environment that must be acted upon. Depending on the cognitive context and the situation, the arousal is subsequently evaluated as positive or negative. Two important differences must be noted between Mandler's theory and ours. First, the basis for the generated expectations in Mandler's view are cognitive schemata. We choose the broader term "predictions" because it can be used for both sensory predictions (in predictive coding) and conceptual (high-level) predictions based on cognitive schemata. Consequently, the term prediction is also more neutral with regard to the conscious access to the predictions. Second, we described any discrepancy (prediction error) as negatively valenced, while for Mandler discrepancies generate undifferentiated arousal, which is later interpreted as either positive or negative. The same discrepancy may produce differently valenced emotions depending on the circumstances and the cognitive context. Instead, we argue that any discrepancy (surprise) is initially experienced as negative, even when a situation is not accurately predicted but in fact better than expected. Instantly afterward, a reappraisal will take place, converting this into a positive experience and actually more positive than when it would not have been preceded by a negative one. This contrast effect will be made more explicit in our discussions of visual art.

In general, the pleasure associated with visual configurations seems to be dependent on the perceptual and emotional dynamics involved. In a similar view, Kubovy defines the pleasures of the mind as "collections of emotions distributed over time whose global evaluation depends on the intensity of the peak emotion and favorability of the end" (Kubovy 1999, p 138). Transitions, rather than static states of stimulation, are needed for these kinds of pleasures. Applied to our context of predictions, positive emotions are experienced when

we have succeeded in reinstating predictability (solving the prediction error). In slightly provocative terms we could say that resistance (of prediction errors) breeds liking.

Another way of thinking about this is that gains in efficiency, the sparing use of resources, are rewarded. If the visual system manages to find a sparse explanation of previously unpredictable stimuli, this genuinely appears to feel good. Similar thoughts have been expressed in the literature on processing fluency (Reber et al 2004). According to these authors, stimuli are preferred more if they are processed more easily. Because of this, familiar, symmetrical, clear-cut (high-contrast), or prototypical (average) stimuli are thought to be liked most. The mere exposure effect (Zajonc 1968) is cited as evidence for this position, seeing that improved processing of a stimulus because of repeated presentations leads to increased preference for this stimulus. At least in case of simple stimuli and subliminal exposures, the mere exposure effect has ample empirical support (Bornstein 1989). If we translate this into the terms of the predictive coding framework, an increase in processing fluency amounts to an increase in predictability (reduced prediction error), but the fluency account has largely disregarded the dynamics of fluency. Recently, however, some processing fluency theorists acknowledged the importance of this in the context of the research on the insight experience (the Aha-Erlebnis). Topolinski and Reber (2010) report that a surprising gain in fluency increases positive affect and the judged truth of the solution found. Since this positive affect sets in before any (conscious) assessment of the "insightful" solution has taken place, they argue this positive effect cannot be solely due to the positive feeling of pride.

Interestingly, discovering the solution of a neutral two-tone image (cf Gregory's camouflaged dalmatian) is accompanied by amygdala activity, and the strength of this activity predicts long-term memory for the organized stimulus (Ludmer et al 2011). Although the study referred to did not directly probe possible associated (positive) affect, it is well established that the amygdala intervenes in reward processing just as well as in fear processing (Baxter and Murray 2002; Sander et al 2003). Here, the amygdala might signal the goodness of the solution found in terms of representational efficiency, and phenomenologically this may be experienced as positive. A reduction of uncertainty, or equivalently, a gain in predictability, might account for the positive experience of the Aha-Erlebnis. That perceptual insight, irrespective of the emotionality of stimuli involved as such, is processed by the same neural structures as those responding to the beneficial biological value of stimuli (Baxter and Murray 2002) is intriguing. It points to the importance of this kind of perceptual and neural (re)organization in economizing representations.

So far, we infer that a (temporary) state of unpredictability (prediction error) can be as important for the emergence of perceptual pleasure as is the predictability. This might particularly be the case for artworks, in which the presence of discrepancies would otherwise be difficult to explain. Earlier (Van de Cruys and Wagemans in press) we used the example of earworms: the melodies that keep haunting us. They may be great Gestalts (extreme predictability), but they are rarely described as beautiful. We will delve deeper into the role of (un)predictability later on, using artworks as examples.

More neural evidence that (un)predictability is an important factor in emotion has been accumulating. For example, a recent study found sustained amygdala activity for temporally unpredictable tones in humans and mice in comparison with rhythmic (predictable) tones (Herry et al 2007). Moreover, both humans and mice show increased anxiety-related responses in a standard anxiety test when these unpredictable (versus predictable) tones were played in the background. Other subcortical areas also seem to be involved in the processing of prediction errors. Research on reward processing has found that dopamine neurons in the ventral tegmental area encode the deviations between actual received rewards

and the expected reward. The reward prediction errors may be crucial signals for learning about rewards (Schultz et al 1997). Another region implicated in reward processing, the putamen, was reported to be more active whenever a tone was not followed by the stimulus type it predicted in an experiment only using entirely neutral, unrewarded stimuli (den Ouden et al 2010). Similarly, plasticity in the amygdala in the context of fear learning seems to be driven by prediction errors originating in the peri-aquaductal gray (McNally et al 2011). Kapp et al (1992) show that amygdala responses wane once contingencies have been fully learned—in other words once stimuli are perfectly predicted—and are reinstated only when external associations change (when predictions fail).

The studies reviewed so far point to an interesting stance, namely that predictions and prediction errors are omnipresent in the brain as has been suggested by several scholars (Bar 2007; Friston 2010a). We have come across predictions in perception—in reward and fear processing and in cognitive schemata. In social and developmental psychology we encounter a long tradition of theorizing about cognitive schemata giving rise to expectations and about the general tendency to reduce dissonance (see, for example, Kagan 2002; Proulx et al 2010), which can also be reconceptualized in terms of predictions and errors. Although a rigorous comparison of these concepts still awaits to be done, we can discern a common theme in all of these studies. The brain's main trade is to predict impending circumstances based on prior, similar experiences. Cases in which this goes awry are significant and therefore emphasized by emotion. The organism has to deal with these mismatches by action (assimilation) or by updating its mental model (accommodation through learning). To show that predictive coding is a pervasive mechanism of mental function might seem of lesser importance for the current purposes. Yet, as an explanation for emotion in art, which evidently works on different perceptual and non-perceptual levels, a theory that encompasses more than purely perceptual principles has substantial appeal.

Finally, this line of reasoning provides a way for visual processing to influence (and be influenced by) emotion. Duncan and Barrett (2007) already remarked that in brain anatomy the boundaries between emotion and cognition seem to dissolve. Conceiving perception in terms of predictions and mismatches opens up a path for perceptual configurations to induce different sequences of affect, partly independent of the particular contents of perception. It is this link we want to explore with aesthetic appreciation of visual art as a concrete application.

Note that in what follows we do not aim to explain all possible emotions in art (if indeed such a thing would ever be possible). We focus on perceptual features of art and only briefly touch upon content, which can, of course, induce a whole range of emotions, from very straightforward ones (as, for example, in paintings of nudes) to very complex ones (as in paintings portraying social settings). Here, we are not interested in intrinsic or associative affective value of the content as such, although we agree that real impressive and expressive art embodies an important interaction between stylistic aspects (perceptual) and thematic ones (see below). Some other affects that we will not talk about but that can show up when contemplating works of art have been listed by Jackendoff and Lerdahl (2006)—for example, the admiration of craftsmanship, emotions linked to nostalgic memories associated with a particular work, social emotions of belonging to (or differentiating oneself from) a group, etc. These different affects probably interact when viewing art, and in the process the sources may be lost, which means that affect caused by one factor can enhance or diminish affect caused by another. Granted that these different emotions are difficult to disentangle in our actual experiences with art, here we are concerned only with the ways in which the perceptual or cognitive organization can provoke emotions.

4 An application to visual art

Visual art is in many ways different from the visual input we ordinarily receive from our natural environment. We want to focus on the prediction errors, more often called incongruities, ubiquitous in art. When we talk of incongruities in the context of art, surrealistic paintings spring to mind. Or perhaps Duchamps' *La Fontaine*, a porcelain urinal, an object one would never have expected to belong in an art exhibition, let alone to be called "beautiful". However, much more subtle kinds of incongruities play in many other styles as well. As a first example, take the painting below by Morandi (Figure 1). In his still lifes he often used subtle differences in hues or texture for figure and background, which makes the perception very unstable. The prediction error resides in those dissolving boundaries and more generally in the fact that Morandi twists expectations of how still-life paintings are defined traditionally.
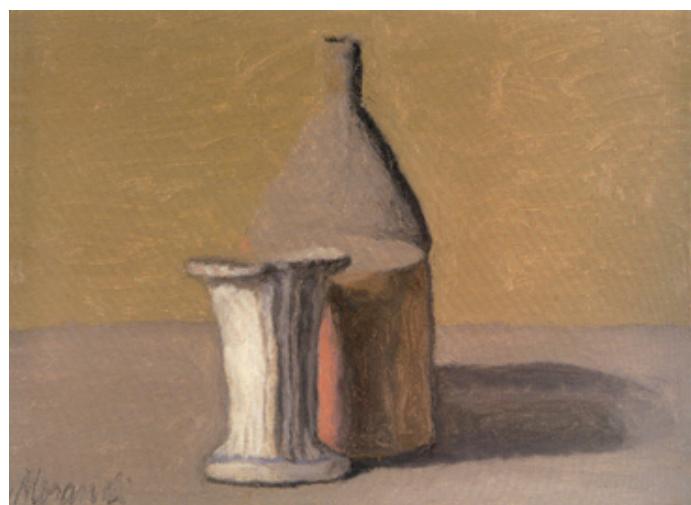
Figure 1. *Natura morta* (1960), Giorgio Morandi.

Why do painters repeatedly and deliberately create such obstacles for us? As we saw, current theories of predictive coding in visual perception assume that our brain aims at reducing prediction errors. In art, however, even though we (or rather, the artists) fully control the stimuli, we intentionally create prediction errors that may not even be possible in the natural visual environment. So in art we do not always manipulate stimuli in ways that reduce prediction errors. Why so? And why does this not imply that we experience most art as unpleasant, as would be the case if all prediction error was negative? The common answer is that in going to a museum, we expect the unexpected. In this limited domain and time span, we can tolerate and even enjoy unpredictability because we expect to be surprised in every new exhibition hall. Furthermore, in these settings there is no need for urgent actions. Prediction errors in art do not need to be acted upon as in real life. This playful and safe as-if context of art, where our guards can be lowered and our actions suspended, allows for the usually negative prediction errors to be enjoyed. Hence, a positive reappraisal can immediately follow the negative gut reaction. But why take this detour? The frequent use of prediction errors in art and their relation to appreciation deserve to be explicated more thoroughly.

4.1 From auditory to visual pleasures

We propose that while prediction error is always annoying or unpleasant initially and confirmed predictions are pleasurable as such (mostly independent of their content), prediction errors or delayed prediction confirmation can be an important tool for artists to amplify the subsequent positive affect of prediction confirmation, in a sort of contrast effect

(Huron 2006). For developing the current view, we are much indebted to Leonard Meyer and his treatment of emotion in music (Meyer 1961). He already invoked expectancy violations to explain emotion expressed in music. Tensions are created in music by first establishing a strong pattern in rhythm and melody (a scale based on the tonic) and subsequently deviating from it. The role of expectancy violations might have been realized in music first because musical patterns are spread out in time, so the dynamics of predictions and deviations thereof are easier to notice. However, they also play in visual perception, albeit on a millisecond time scale. For example in *Day and Night* by Escher (Figure 2), every new bird-like form the visual system encounters confirms and strengthens the (later destroyed) parsimonious prediction.
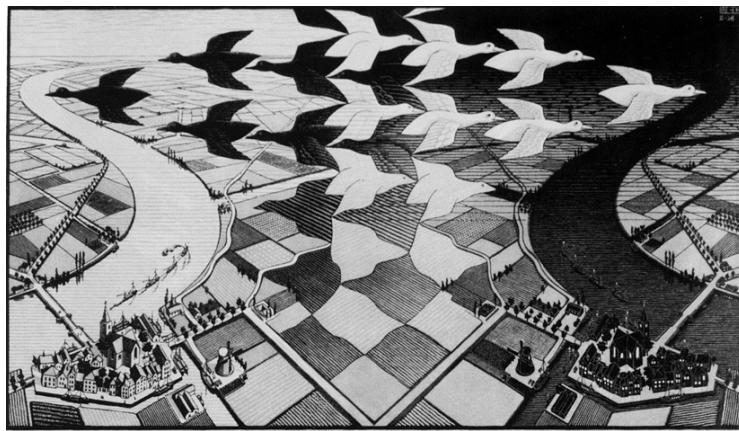

Figure 2. *Day and Night* (1938), M C Escher.

An experience of deep aesthetic appreciation is not so easily reproduced in the lab. One more reason music might be more tractable in the lab than visual art is that music happens to be the art form in which an aesthetic experience can be best elicited in an intense, more or less controlled and reliable (repeatable) fashion. These experiences known as music "chills" or "shivers-down-the-spine" have been successfully used to explore the neural and cognitive underpinnings of music appreciation (Blood and Zatorre 2001; Huron 2006). Cognitively, Sloboda (1991) found that these chills are strongly correlated with marked violations of expectation. Neurally, Blood and Zatorre (2001) reported that they are associated with increases in cerebral blood flow in regions involved in reward processing (ventral striatum) and decreases in amygdala and ventromedial prefrontal cortex. Similar to the study by Ludmer et al (2011) discussed earlier, changes in predictability of stimuli elicit a pattern of activity characteristic of processing biologically relevant, survival-related stimuli.

While musicians use uncertainty on the when (rhythm) and the what (melody, volume) to evoke emotions in their listeners (Huron 2006), visual artists utilize uncertainty on the where (spatial) and the what (object indeterminacy, see below) with the same intent. Below we will discuss a series of examples, elucidating the workings of this artistic "tool".

In *Separation* by Munch (Figure 3) there is a (mild) violation of grouping by similarity of color and form; however, we have no problems finding out what the objects are. Artists seem to like using strong predictions, either by building them themselves through repetition, as in the Escher painting above, or by using a well-known conceptual domain in their painting and thereby relying on existing strong predictions. In Munch's *Separation* we find

this combination of a familiar pattern (strong predictions) together with a minimal deviation of default expectations, which makes the painting attention-grabbing and memorable.[1]

Figure 3. *Separation* (1896), Edvard Munch.

In his *Weeping Woman* (Figure 4) Picasso arguably counts on our specialized face-processing systems (fusiform face area) to project its guesses on what belongs where in a face based only on some fragmentary cues. The viewer quickly runs into incongruencies, which presumably generate arousal aimed at reducing the prediction errors. This style-induced arousal could add up the emotionality of the contents, because despite the "errors," we can still recognize the emotional expression portrayed. We can contrast this with Ramachandran and Hirstein's (1999) explanation invoking the peak shift phenomenon. In this account, Picasso's face derives its emotionality from being a superstimulus, activating visual face processing systems very strongly because input from multiple viewpoints is merged in one image. This explanation for art is refuted by Hyman (2010), who observes that artistic depictions often deviate too far from the norm to be examples of peak shift (which is, as Hyman reviews, a very specific phenomenon in animal behavior). In our view it is precisely the incompatibility (prediction error) that causes part of the emotionality in this (and other) paintings.

Predictive coding is an intrinsically hierarchical, multi-level model (Lee and Mumford 2003) in the sense that implicit predictions are generated and checked on every level in the visual system, from low-level feature-related to mid-level configurational predictions, up to high-level concrete and abstract semantic predictions. Importantly, predictions on every level are partly encapsulated (Jackendoff and Lerdahl 2006), such that even when we are very familiar with a painting (eg, *Blanc Seing* by Magritte [Figure 5]), and the visual input is in fact not unexpected any more, we are subject to these dynamics. That is, our visual system unconsciously computes its moment-to-moment predictions and errors regardless of the viewer's conscious memory.

---

[1]Of course, there is more to the paintings we discuss than the aspects we zoom in on. We just want to draw attention to a possible commonality in these pieces, which can appeal to both laymen and experts.
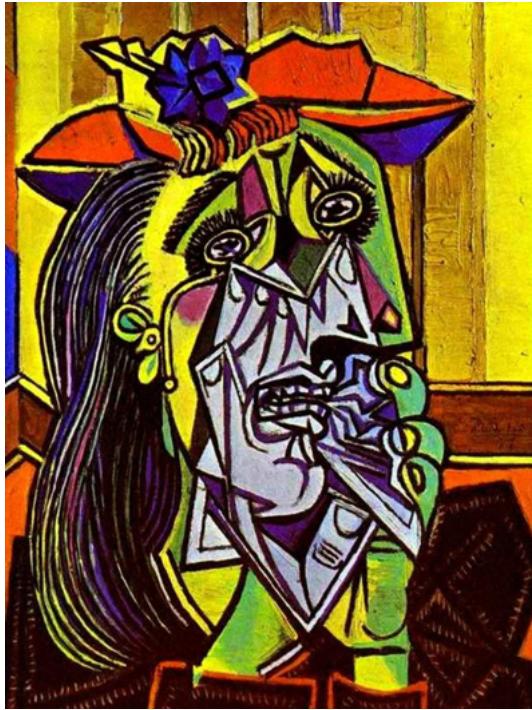
Figure 4. Weeping Woman (1937), Pablo Picasso.
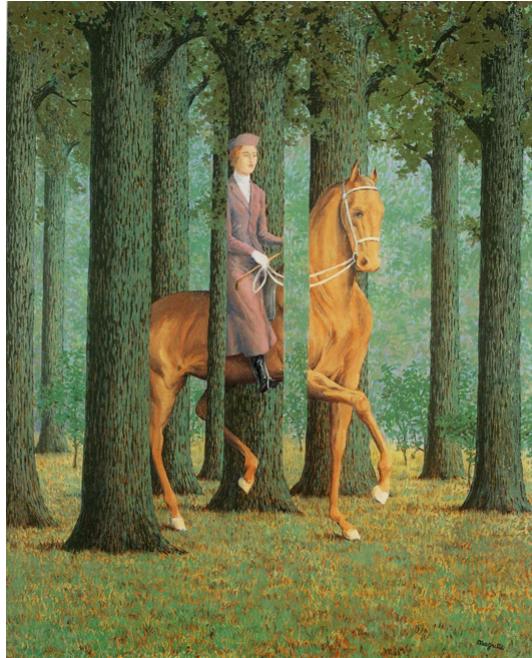
Figure 5. *Blanc Seing* (1965), René Magritte.

4.2 Art movements

Vincent van Gogh in *The Olive Trees* (Figure 6) plays with perceptual grouping by similarity (parallel waves) breaching the borders of the objects as defined by color and by our top-down knowledge of what the objects are in the scene (trees, fields, sky). In this light, differences between art movements and artists might also be interpreted in terms of differences in the amount and kind of prediction errors primarily used. The evolution of movements can be understood within our framework, recalling that predictions are dependent on the

specific history of stimulation, and therefore on cultural and personal experience. Recall that van Gogh (and numerous other great artists) did not enjoy any recognition from his contemporaries. He was mocked and died in obscurity. Predictions evolve, and what was a prediction error once can now be the canonical form of expression and thus fully predictable. These shifts partly explain why artistic taste varies widely in different eras and within the same era, between experts and laymen. In the predictive coding approach all perception is a form of expert perception (Clark 2011), in the sense that it is always determined by an individual's expectations built up through a lifetime of implicit statistical learning.
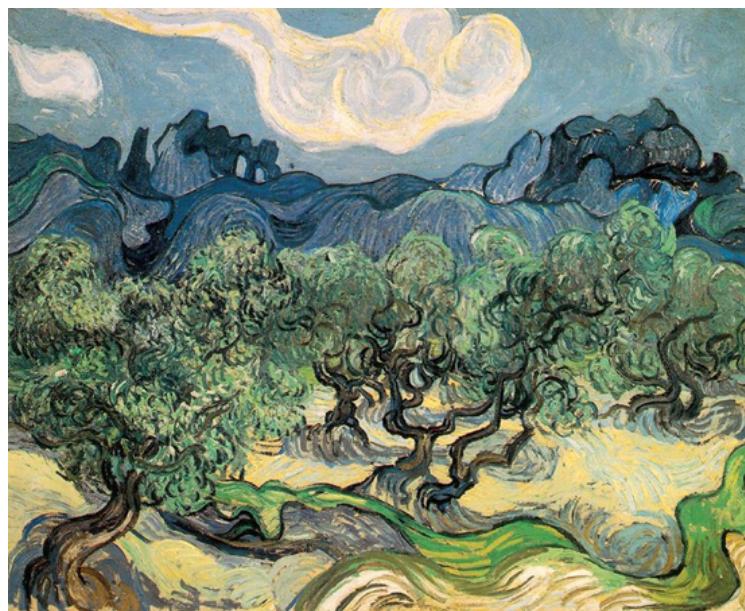

Figure 6. *The Olive Trees* (1889), Vincent van Gogh.

A certain kind or amount of prediction error can be upgraded intentionally or unintentionally to being the norm, the predictable standard, only for those who have developed the allegedly exquisite sensitivities and expertise needed to appreciate or grasp these artworks (ie, seeing the underlying predictable structure). This role of aesthetic taste in creating social distinction and status is defended by the sociologist Bourdieu (1984), who argues that by openly declaring oneself an aficionado of difficult and inaccessible artworks people reassert their membership in society's upper classes. It also figures in current theories on the evolutionary role of art as an expensive and time-consuming status symbol, not unlike the peacock's tail (Pinker 2003).

4.3 Inter-individual differences and the optimum of (un)predictability

We can derive two important hypotheses from the previous. First, it follows that individuals are likely to have an optimal amount of unpredictability that they most appreciate. Too much prediction error is unpleasant or even disturbing; none or too little is boring (neither positive nor negative). This relates to psychological or subjective complexity, an important determinant of people's aesthetic evaluations (Gaver and Mandler 1987). The psychological complexity is a function of both actual stimulus complexity and personal experience with the class of stimuli involved. The predictability of the work of art increases with the observer's knowledge of the correlational structure built up through exposure. An optimum of mild violation of predictions will be experienced as most pleasant, because the experiencer manages to return to a familiar mental schema. Therefore, for a given piece of art, we will also see an evolution of liking through repeated exposure, with initially an increase of preference

and motivation to enjoy the piece time and again, but as predictability further increases appreciation will diminish again. A similar evolution of preference has been reported by Berlyne (1970) relating complexity and novelty to arousal. He hypothesized that stimuli of moderate complexity engender a moderate amount of arousal, which is optimally pleasing. Berlyne does not further address the question of whether this mild arousal can be pleasant in itself or, alternatively, whether the moderate complexity still allows a resolution to a predictable, recognized configuration, which only then is experienced as positive. We will return to this issue later on.

Reliable inter-individual differences in appreciation should thus be present, which will be dependent on personal experience, which is especially obvious when comparing aesthetic evaluations by experts with those of laymen (Lindell and Mueller 2011). Additionally, we speculate that these inter-individual differences in preference also depend on the strength of the top-down predictions the viewer generates (irrespective of the particular perceptual predictions). This factor may determine how much prediction error a particular person can "tolerate" and appreciate, but whether this is the case is still an open, empirical question. Related to that, one may wonder whether personality factors such as personal need for structure (Thompson et al 2001), field dependency (Witkin and Goodenough 1981), or the systemizing quotient from autism research (Baron-Cohen et al 2003) relate to this strength of top-down predictions. At least for personal need for structure a correlation has already been found with art appreciation (Landau et al 2006).

The second hypothesis derived from our theorizing is that artists attempt to strike this optimal balance between predictability and surprise in their works. This way their viewers have to make an effort and initially experience minor negative affect, only to experience a much intenser positive affect by contrast, once they actually mentally "resolved" the prediction error. The effort or mental work one has to do to cope with the prediction error is a conditio sine qua non for receiving perceptual pleasure of a Gestalt formation (prediction error reduction). Before arriving at a clear, coherent interpretation of the jumble of lines in Klimt's *Reclining Woman* (Figure 7), our visual system is embroiled in a small struggle. Only with some mental work is the familiar silhouette discovered. But the aesthetic pleasure is larger as a result. Sometimes this Aha-Erlebnis emerges only after a glance at the title for some contextual information, as for Picasso's *Guernica*.
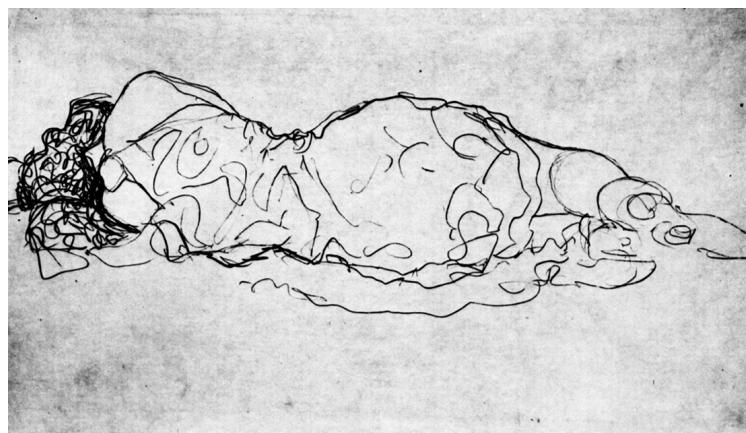

Figure 7. *Reclining Woman* (1914/17), Gustav Klimt.

The hypothesis that aesthetic pleasure lies in forcing the brain to do some work has also been put forward by Dodgson (2009). His research starts from the computer-generated geometric patterns inspired by the early work of the English Op artist Bridget Riley. With

a psychophysical test using different degrees of distortion, he shows that there is a range in which a pattern is not immediately recognized but can be recognized given some extra effort. Since Riley's works precisely lie in this region where a pattern is hinted at but not made obvious, Dodgson suggests that this might be a common mechanism in all art. In line with our account above, he argues that finding an aesthetic optimum is not a matter of depicting the best Gestalt but rather of providing just enough information that the viewer can reconstruct the pattern but not so much information that the pattern is plainly obvious. Hence, Dodgson's research is a nice illustration of the idea that artworks embody a right balance of the expected and the surprising. It also helps to clarify discussions on the relative importance of familiarity and novelty in art appreciation where it is often concluded that: "It is not extreme novelty but 'optimal' innovation—novelty that allows for the recoverability of the familiar—that is most pleasurable" (Giora [2003, p 176] quoted in Lindell and Mueller 2011, p 465).
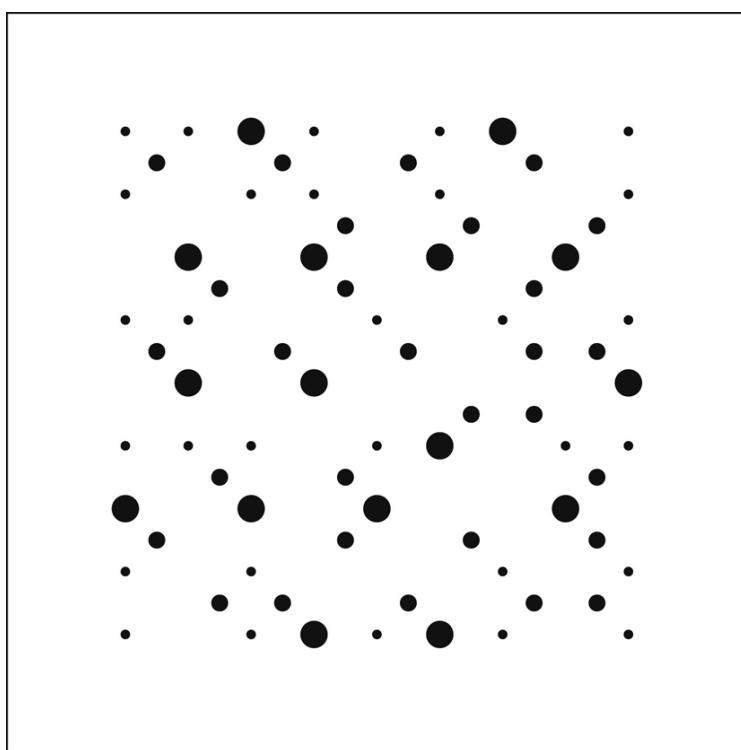

Figure 8. Computer rendition of Fragment 6/9, Riley (from Dodgson [2009] with permission from the author).

4.4 The valence of prediction error

A central assumption of our theory is that prediction errors are always to some extent emotional, more specifically negative in valence. We are currently setting up experiments to test this, but so far we have only indirect evidence to support it. For example, we note that the expression of surprise, the emotional reaction to violation of predictions, is very close to (and often indistinguishable from) a fearful expression. Neuroimaging studies lend support for this similarity since the amygdala responds equally strongly to surprised as to fearful faces and much less to angry or happy emotions (Whalen 1998). Moreover, a very recent meta-analysis of neuroimaging studies on aesthetic appreciation (across modalities) found a consistent activation of the anterior insula (Brown et al 2011). This structure is known to be involved in the conscious, bodily emotional experience (Craig 2009), in particular for unpleasant emotions such as disgust and pain. The authors were puzzled that this region

systematically pops up in studies on aesthetics and assumed it attests to the visceral impact of artworks. In light of our theory, the insula activity might point to the unpleasantness of prediction errors, which are an important stage in the aesthetic experience. The limited temporal resolution of fMRI might not make the distinction between stages that quickly follow each other. However, we should tread carefully in interpreting the anterior insula activity because, as Yarkoni et al (2011) report in their large-scale meta-analysis, the insula is one of the regions found to be active in a large percentage of neuroimaging studies, a lot of which do not even study emotion. Still, it is interesting to see that the anterior insula is not only involved in negative emotions but that more specifically (spatially and temporally) separable signals can be found reflecting risk prediction and risk prediction errors in the insula (Preuschoff et al 2008). This role of the insula in uncertainty and anticipation, which has been noted before (Elliot et al 2000), is in line with our reasoning that prediction errors in art cause uncertainty (risk of not being rewarded in the end), the computations of which will determine whether to look away or to go on exploring the piece. Moreover, the anterior cingulate cortex, a region often jointly active with the insula (Craig 2009), has been found to be involved in signaling visual prediction errors (Noyce and Sekuler 2011) analogous to its role in signaling motor mistakes (which we can reasonably assume are experienced as negative). All this being said, whether our hypothesis of the negatively valenced prediction errors stands the empirical test remains to be seen.

While traditionally the aesthetic experience is equated with attaining harmony and successful insights, we are not the first to emphasize negative affect, discrepancy, and disruption as important components in aesthetic appreciation. Vygotsky (1971) used Aristoteles' concept of catharsis to describe that in aesthetic experiences we go through a transformation from a negatively valenced obstruction to a positively experienced resolution, in full awareness that the positive catharsis would not be possible without first having experienced its contrast. More recently the artist Pepperell (2011) characterized his aesthetic experience as a moment in which his "usual conceptual grip on the world failed". This was accompanied by mild anxiety and "active struggle to make sense of what [he] was seeing". Artists use the compulsive interpretative (predictive) mode of the visual system to leave out parts of, or blatantly contradict, the suggested content of their works.

Pelowski and Akiba (2011) similarly highlight disruptions and breaches as vehicles for the kinds of self-transformation described in accounts of rich aesthetic experiences. When artworks forcefully challenge the conceptual classifications and personal self-understandings of the viewer, they can lead him or her to question, change, or expand these, with possible existential ramifications. This is a precarious balance to strike because the viewer might prematurely break off perception. Overcoming initial negative experiences requires courage and inventiveness to modify the existing (self-)schema (Vygotsky 1971; Pelowski and Akiba 2011) but is constitutive for the full aesthetic experience. It is encouraging to see that, starting from different inspiration, namely the full, rich aesthetic experience and several social-psychological findings, the authors arrive at an understanding of the aesthetic experience that is consistent with ours. While their view explores the existential and meta-cognitive implications of the conflicts present in art, ours starts from a substantiated theory of visual perception, in a rather "bottom-up" way.

4.5 No pain, no gain?

What evidence do we have to argue that a reduction in unpredictability (prediction error) is experienced as positive? Put differently, do we have reason to assume that the degree of our mental efforts to compensate for this unpredictability (to resolve the picture) is directly related to the reward we experience? A common finding in animal learning tasks is that random intermittent rewards (there is a pattern but it is disrupted) or non-reward (during

initial extinction trials: there was a clear pattern, but now it is interrupted) provoke most vigorous responding. In those cases motivation is highest. For example, a variable interval reinforcement schedule, in which the animal is rewarded on average every n^{th} amount of time, provokes the highest response rates, presumably because from the point of view of the animal there remains something to be learnt. Neurally, unexpected rewards, compared with expected rewards, are associated with increased dopamine peaks. Johnson and Gallagher (2010) specifically asked whether it might be the effort required for getting a reward that is the determining factor for the reward value. They reported that the positive affective quality of taste can be boosted by increasing the amount of effort required to obtain it. In their study mice learned to press a lever A for a reward and lever B for another, different tasting, reward. Next, they gradually increased the number of operant lever presses needed to obtain one of the two rewards. When tested afterwards outside of the training environment, the mice showed a clear preference for the reinforcer for which they had to work the hardest in the learning phase. In accordance with our speculations on art, they concluded that pleasure (hedonic value) is increased by effort. In humans something similar has very recently been reported as the IKEA effect: participants value the products of their own labour similar to creations of experts, but only when they successfully completed them (Norton et al in press).

The view outlined here is bound to learn a lot more from advances in (animal) learning studies on the link between uncertainty versus predictability and reward value. They might help to explain why we can derive pleasures from mildly deviant stimuli, as in an addict who experiences pleasure from almost winning (near-misses). In a set-up mimicking a gambling machine, almost winning is shown to induce heightened reward expectancies in rats, mediated by dopamine and analogous to the near-miss effect in addiction (Winstanley et al 2011). It seems that part of the reward lies in the anticipation instead of the consumption (Lauwereyns 2010), with a little unpredictability adding to the reward. One must note that this is in agreement with the idea that biological organisms are characterized by striving: It would have made little sense for evolution to reward static states of stimulation. Prediction errors introduce uncertainty that will stimulate further processing (mental effort), the outcome of which is also uncertain, thus unexpected when successful. Phenomenally, this kind of perceptual problem solving is not only rewarding but also gives the viewer a sense of mastery (Leder et al 2004).

In the context of art perception, Ishai et al (2007) find a positive correlation between the time needed to comprehend a picture and its aesthetic value (measured as the judged degree of "powerfulness" of a presented painting), leading Pepperell (2011) to suggest that the effort people have to invest to recognize the contents has a positive influence on aesthetic value. Reports that the better the viewer's understanding of an artwork, the more he or she experiences pleasure, should thus at least be re-examined (Leder et al 2004). No doubt individual differences determine a large part of the variability here (as described earlier), but in addition to that, experiment designs may not always allow the tracking of these dynamics and the presence of moments of misconception. It may be true that understanding facilitates liking, but this pleasure may be potentiated by preceding mismatches. Likewise, when going to a scary film, we endure cruel tensions because their resolution is so good. Note, however, that when we say we like a painting, there is a kind of misattribution taking place: the pleasurable experience associated with a processing characteristic, namely the reduction of prediction error in the stimulus, is ascribed to the stimulus itself (Huron 2006).

Our fundamental thesis is that the progression from a state of more unpredictability to a state of less uncertainty is pleasurable. Again, animal studies support this hypothesis. In an elegant experiment Bromberg-Martin and Hikosaka (2011) showed that monkeys have a preference for cues that indicate that advance information on the presence or absence of

a forthcoming reward will be available over cues that signal such information will not be available, even if their choice has no influence whatsoever on the actual appearance of the reward. Moreover, the monkeys more consistently chose this advance information over the no information cue than they choose a more probable reward over a less probable reward (Niv and Chan 2011). In other words, information or reduction of uncertainty is rewarding as such, consistent with our account. In fact, when Bromberg-Martin and Hikosaka (2011) examined the neural underpinnings of the value of information, they concluded that it is encoded by the same habenula and dopaminergic neurons that also encode primary rewards. To be clear, these dopamine and habenula neurons are known to signal a change in reward contingencies (reward prediction errors), and this is also what they found with regard to information: these neurons do not encode the value of predicted information per se but only the changes in predicted information (Niv and Chan 2011). The signals related to information change were simply added up to signals related to reward changes, indicating that the value of reward and that of information use a interchangeable neural currency (Niv and Chan 2011). In keeping with this, we contend that decreases of (information) prediction error are pleasurable and constitute part of the aesthetic pleasure. Cues that signify a reduction of uncertainty also seem to have greater value for humans. Ogawa and Watanabe (2011) asked people to perform a contextual cuing task in which targets are surrounded either by a repeated (and thus predictive) configuration of distractors or, in the other half of the trials, by a novel display of distractors. When subjects had to evaluate the goodness of the displays afterward, predictive displays were judged more positively than non-predictive or novel configurations. Since people did not see predictive displays more frequently (nor did they recognize them significantly more in a final phase of the experiment), this increased liking could not be caused by mere exposure. The authors conclude that predictability promotes preference (Ogawa and Watanabe 2011). In the context of an effortful search task, with a large degree of uncertainty on the position of the target, predictable configurations will become associated with a higher value than configurations that do not reduce uncertainty.

To sum up, dealing with unpredictability requires effort from the viewer. But when successful, it leads to positive appreciation. The idea that the viewer's personal efforts and "accomplishments" matter in art appreciation is also recognized by Mamassian (2008, p 2152): "not all ambiguities in paintings are re-solved, and artists probably strive to leave the right amount of ambiguities to let the observer contribute to his experience in a personal way." However, only by using minimal prediction errors painters can ensure that viewers will obtain their reward(s) and not give up prematurely. Final gratification may be further postponed as long as the artist has hidden in the painting enough "micro-rewards" the viewer can discover along the way.

4.6 Dynamics in art

Our view emphasizes the role of the dynamics of perceptual processing in art appreciation. This emphasis is not particularly new, but has been foreshadowed by Arnheim, although arguably in a less articulate way. In his seminal work, Arnheim (1974) asserts that visual experience is dynamically governed by "attractions" and "forces" or "tensions". Even though these terms are kept rather vague in his writings, our prediction error account could be seen as an effort to give his concept of tension new substance. Arnheim realized that these tensions are inherent in any percept, as with our concept of prediction (error), even acknowledging the active role of the observer and his past experience in (automatically) generating these tensions. Analogously to our view, Arnheim contends that psychological, like physical, systems "exhibit a very general tendency to change in the direction of the lowest attainable tension level" (Arnheim 1974, p 14). A crucial difference between Arnheim's view and ours is that the ultimate goal for Arnheim is balance or stability, while for us it is the reduction

of prediction error (or maximization of predictability). In general, once he acknowledged the dynamic nature of perception, he remains overly focused on balance and stability as the static pinnacle of beauty, largely ignoring the emotional dynamics of the transitions going from "tensed" to "relaxed" interpretations of visual input that, according to our view, may cause positive appreciations.

Because it is inspired by predictive coding, our view also shares with Arnheim's the concern for parsimony (efficiency of representation, cf supra), which he in turn inherits from the Gestalt tradition, with its emphasis on simplicity and *Prägnanz*. Predictability implies more efficient, higher level representations, freeing up expensive processing resources. That there is positive affect (appreciation) associated with this progression has been anticipated by Eysenck (1942) in his law of aesthetic appreciation: "The pleasure derived from a percept as such is directly proportional to the decrease of energy capable of doing work in the total nervous system, as compared with the original state of the whole system" (Eysenck 1942, p 358). Thus, a change towards a more efficient state of the dynamic perceptual networks given the current constraints of stimulation, tantamount to a decrease in prediction error (surprise), is pleasurable as such, as we described for the Klimt sketch above (Figure 7).

4.7 From traditional to modern art

Melcher and Bacci (2008) noted that artists often use very familiar perceptual domains, like faces or animals, but apply their "artistry" on them. Since prehistoric times, artists exploit very familiar patterns, for which we developed exquisite, specialized systems through evolution and experience. Melcher and Bacci (2008, p 352) observe that "the ability to recognize certain stimuli quickly and easily makes it easier, then, for artists to add decorative elements, accurate details and artistic style." For example in prehistoric cave art, the patterns portrayed are not "just" depictions of animals or humans; they incorporate counter-intuitive or unusual deformations (see Figure 9).
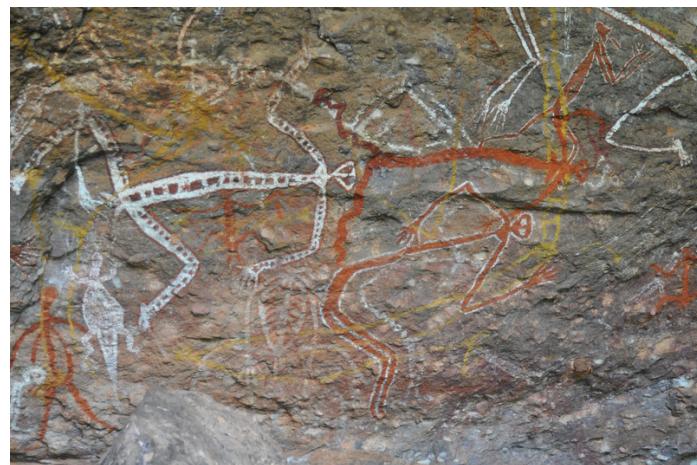
Figure 9. Examples of prehistoric cave art (elongated human figures in aboriginal art in Kukadu National Park, Australia).

The cultural anthropologist Dissanayake (2009) calls this special treatment "artification" or "making special," which, according to her, encompasses formalization, exaggeration, elaboration, repetition, and manipulation of expectation. Note that most (if not all) of these terms relate to either confirmation and strengthening of predictions or violation and disruption of predictions (prediction errors). Dissanayake agrees that this process is most obvious in the temporal, performing and traditional arts, but claims that these are universal characteristics common to all art forms. This particular shaping of stimuli is meant to set

these objects or practices apart from, and make them more than, the ordinary (traditionally in the context of religious rituals).

In the context of their discussion of surprise in art, Melcher and Bacci (2008) provide some other fine examples of prediction error in art. About the *Sleeping Hermaphrodite* they write:

> Displayed with the back to the viewer, it offers a sensual pose that flatters the femininity of the figure, in accordance with the tradition of the reclining Hellenistic nude. Enticed by the sensual nature of this work, the viewer who will then walk around the sculpture to see it from the front will encounter the surprise of the figure's androgynous nature, which provides a moment of astonishment in which sensory information does not coincide with expectation (Melcher and Bacci 2008, p 354).

On the powerful, realistic evocation of biological motion in Caravaggio's *The Death of the Virgin* they hypothesize: "Thus, perception of motion in static art involves a form of perceptual ambiguity, in which complex motion areas such as MT/MST are activated but 'early' vision detectors in V1 are silent" (Melcher and Bacci 2008, p 357).

Several theorists and art critics have argued that (post)modern artists, with their explicit abandonment of any representational aspirations and their blunt refusal to provide any familiar reference point whatsoever to their spectators, have taken things too far (Landau et al 2006; Pinker 2003). They are said to leave the viewers of their works completely in chaos, bombarding them with colors and fragmented shapes with no structure or meaning. In our terms: an overload of prediction errors that despite our best efforts, cannot be reduced to a more predictable, sparser explanation. As a consequence, modern abstract art, as compared with other styles, is least preferred by the general public (Landau et al 2006; Lindell and Mueller 2011). We suggested earlier that artists will attempt to find an optimum, but their success in doing so depends on their intended audience and its broader cultural context. Works of art will survive through time if they manage to be at such an optimum (and thereby appeal) across cultures and periods (Melcher and Bacci 2008). Of all abstract art our theory predicts that we will be most impressed by paintings in which we can, with some labor, distinguish recognizable, predictable forms. *Excavation* by de Koonig (Figure 10) can serve as a good example, as can the paintings of Pepperell that also show this kind of object indeterminacy, investigated by Fairhall and Ishai (2008).

Figure 10. *Excavation* (1950), Willem de Koonig.

Clearly our approach can be applied to art belonging to a broad range of styles and eras. In fact, it has most difficulties in explaining the attraction of hyper-realistic art. In this case we can speculate that it has more to do with the admiration of craftsmanship, the emotional content of the painting, or possibly an ultimate prediction error in all visual artworks: this work provokes strong feelings in me, as if the particular object depicted is really in front of me, while in fact it isn't. What we could call the "*ceci n'est pas une pipe*" experience. Alternatively, we might be bemused by the fact that we see a nearly photo-realistic image while we expected a painting. But evidently these are just ad hoc explanations.

4.8 The fate of prediction errors

The examples so far might have made it clear that often we do not end up with a coherent, predictable Gestalt. Still, we want to draw attention to the fact that prediction errors at the level of style (perceptual ones) sometimes can be resolved on the level of meaning (see, also, Pinna 2010). The deserting lady in Munch's *Separation* dissolves in the road and the air, and Picasso's *Weeping Woman* can be "broken" because she is sad. A clear congruency in meaning unexpectedly "saves" the coherence of the paintings, and so the prediction error is reinterpreted. With these symbolic explanations, we enter a more speculative realm, but it would be foolish to expect that something as complex as art appreciation can be understood by looking only at the low-level, perceptual features. Art is about the interaction of style and content, not a simple addition of these components. As we reviewed in the introduction, a predictive coding view can, similar to schema theory (Proulx et al 2010), provide a way for high-level expectations to be involved in our experience with art.

What about those paintings in which we can not even remotely and mentally come to a form of closure? Paintings are static art forms, so prediction errors often cannot be resolved, except in our minds. In their artistic endeavors people seem to deliberately seek prediction errors. Prediction errors intrigue us, especially when they violate strong default expectations. Even if we are not sure there is an actual clue to them, we cannot remain indifferent to them (they cause arousal), and we keep coming back to examine them (attentional resources are recruited). We might still experience positive emotions from these "unsolvable" paintings, though. By reappraisal of the negative prediction error in a safe context, the resulting emotion is still very positive because of the contrast effect (cf supra). A related, but more implicit mechanism, namely misattribution, could turn a negative arousal into something of positive affective valence. This mechanism shows people often have very little insight into the sources of their experienced emotions or arousal, as exemplified by the classic Capilano suspension bridge experiment (Dutton and Aron 1974). In this experiment passers-by on a scary rope bridge or a solid wood bridge are asked some sham questions by an attractive female researcher and told that they can always call for more information on the research afterward. The authors found that if participants were questioned on the foot-bridge, they were much more likely to call the researcher afterward to ask her out. The authors assumed that those participants misinterpreted their arousal from the fright of walking on a high, shaky bridge as feelings of attraction. The specific experiment has been criticized, but the phenomenon of misattribution has been replicated since and is well-established. It might thus partly explain why we misinterpret, for example, the mildly negative color mismatch in a Matisse painting as affectively positive (Figure 11).

The question of whether prediction errors can be experienced as positive also relates to the conceptual confusion on what exactly an aesthetic experience entails. We can safely assume that art has to be rewarding in one way or another; otherwise, we would not be motivated to engage with it. But people might label some artworks as "fascinating" or "special", rather than "beautiful" (Augustin et al in press). Finally, a particular painting

Figure 11. *Blue Nude* (1952), Henri Matisse.

might also be differently appreciated depending on the context of stimulation, as we will briefly discuss in the next section.

4.9 Inferentially rich, attention-grabbing meaning threats

Because of the prediction errors, we feel impelled to question our perception and to linger on its contents. These visual or cognitive challenges urge us to, implicitly or explicitly, go through multiple cycles, exploring different predictions and the corresponding errors (Leder et al 2004). They grant access to different layers of meaning, which we so much like to discover. They create the multi-interpretability and ambiguity that has been invoked by others to explain our enjoyment of art (Biederman and Vessel 2006; Mamassian 2008; Zeki 2004; van Leeuwen 2007). In the context of creative discovery, Verstijnen et al (1998) have coined the term *surplus structure* to denote that through externalizing their ideas in sketches, artists themselves discover new, unanticipated features and interpretations of the raw ideas. Might this be true? Not so much because in sketches they can more lucidly represent their ideas but rather because sketches allow them to depart more easily from ordinary ideas and exaggerate, restructure, and deform more freely, in other words amplify what would normally be prediction errors?

Biederman and Vessel's (2006) idea is basically that inferentially rich stimuli will be preferred because they are accompanied by more activity in regions higher up in the ventral visual stream, which possess higher amounts of mu-opioid receptors. This hypothesis is not at all incompatible with our account. Our optimally deviant expected patterns with a limited violation of intuitions, combine the rich inferences (of a predictable pattern) with the high saliency (of the discrepancies) which conspire to make a very emotional and memorable stimulus (Sperber and Hirschfeld 2004). Discrepancies are attention grabbing and stimulate

further processing, but only when strong predictions are first built up (clear organization). This optimum makes for a highly relevant stimulus according to Sperber's (2005) theory because it guarantees the richest cognitive inferences for the least cognitive effort.

Our account also bears a clear resemblance to the Freudian notion of the uncanny (*das Unheimliche*). Freud observed that we are disturbed and aroused by unfamiliar experiences in an otherwise completely familiar setting (Proulx et al 2010). For example, in absurdist or surrealist art we often find an unfamiliar juxtaposition of very familiar objects. Thus, strong expectations have to be present before such an experience can ensue. In a completely unfamiliar situation no strong predictions are formed, so no violations will be encountered. Incongruency or expectancy violation will result in negatively valenced arousal aimed at reducing the inconsistency if at all possible, as in our account. Some authors have even reported that this arousal can lead to an affirmation of any other meaning framework to which one is committed. For example, after exposure to an absurd Kafka parable, subjects more strongly affirmed their cultural identity than after reading one of Aesop's meaningful parables (Proulx et al 2010). The meaning maintenance model, as this theory is called, has been developed to counter the terror management theory, which assumes that only mortality threat (salience) will cause people to defend their cultural world-view. In a study by Landau et al (2006) terror management theory has also been applied to art appreciation. Apparently, mortality salience (having people to imagine what will physically happen when they die and which emotions they have with that) decreases their liking for "meaningless," abstract art, while leaving their appreciation of representational paintings untouched. In further studies the authors find that this effect is limited to individuals with a high personal need for structure, and diminished when the abstract artworks were given meaning (eg, by giving a title). This research goes to show that next to stable traits of the viewer and stable characteristics of the piece of art, aesthetic appreciation can also be influenced by the context-dependent cognitive and emotional mindset the viewer is in. We could speculate that any contextual uncertainty (prediction error or, in terms of the theories discussed, threat of meaninglessness) could add up to the uncertainty in a painting and thus influence its appreciation. For instance, against a background of unpredictable (versus rhythmic, cf supra) tones a representational painting may be liked more, while an abstract painting might be liked less. Related to this, Mueller et al (in press) recently probed implicit and explicit attitudes towards creativity after inducing a sense of uncertainty in half of their volunteers. These people were told they might receive additional payment based on a random lottery or, in a second experiment, they were primed to be intolerant of ambiguity. While explicit attitudes towards creativity were similar in the experimental and the controls, people in the high uncertainty condition had an unconscious bias against creativity and judged a highly creative idea less favorably.

Ultimately, maintaining (or returning to) predictability is about survival and maintaining the body through homeostasis (Cerra and Bingham 1998; Van de Cruys and Wagemans in press; Friston 2010a). Predictive coding is about reinstating predictability and therefore about affirming one's own existence. So within a predictive coding framework, we do not need to assign a special status to the existential threat of mortality. We further assumed that making progress in this prediction project of life genuinely feels good or, in the case of art, is beautiful. In his *Ethics*, Spinoza (2001, p 40) writes, "[I]f the motion by which the nerves are affected by means of objects represented to the eye conduces to well-being, the objects by which it is caused are called beautiful; while those exciting a contrary motion are called deformed." When a successful, sublime aesthetic experience is described as a selfless state of harmony between the viewer and the world (Pelowski and Akiba 2011), we might take this quite literally: the beholder has advanced in tuning the self to the world.

5 Remaining questions and concluding remarks

We hope to have shown that the predictive coding approach can summarize and throw a new light on existing concepts in the flourishing field of aesthetic appreciation, such as familiarity, complexity, novelty, prototypicality, interpretability, fluency, incongruency, ambiguity, and so. Further research will have to make clear what the added value is of thinking in terms of predictions and prediction errors in comparison with these concepts. We want to end our overview by discussing some of the limitations and advantages of our approach.

First of all, we do not want to reduce aesthetic experience to the formal mechanisms discussed. For instance, we are not saying we experience a full-blown aesthetic reaction when discovering the actual content or organization of 2-tone images, but our hypothesis is that a significant part of the aesthetic appreciation comes from discovering the organization after struggle.

Second, we are being too vague when we claim that art is about optimally or minimally unpredictable stimuli. The amount but also the kind of prediction errors seems important. Some prediction errors are more potent than others. For example, we saw that artists either induce strong predictions themselves in their viewer or rely on strong existing predictions of the domain used in the painting, to subsequently violate them. Also, different artists may have different preferences for the kind of prediction errors they use, as reflected in their style. For instance, some artists play with classical grouping principles and their competition.

Third, one might object that our focus on prediction errors is born out of cultural myopia. In Western art there is a strong impetus to be original and novel, and even to defy established traditions. In traditional, non-Western cultures, however, originality is often discouraged, and artists are expected to closely follow and endlessly repeat the same set of patterns passed on for centuries (Dissanayake 2008). As we mentioned earlier, this repetitive art also involves a modulation of predictability but it seems to lack any prediction errors. However, Dissanayake also notes that these forms of art originally take place in the context of ritual ceremonies in times of transition or uncertainty. These rituals nearly always concern biologically important things, such as "assuring or restoring subsistence, safety, fecundity, health, prosperity, and victory or successfully dealing with the bodily changes and emotional and social concomitants of sexual maturity, pregnancy, birth, and death" (Dissanayake 2008, p 19). Hence, she sees stress reduction or coping with uncertainty as an important adaptive function of art. If we assume that, at least in these particular ceremonial situations, people in traditional societies experienced more life-threatening uncertainties than we do in our modern Western society, our hypothesis of a general preferred optimum of unpredictability could still hold. In traditional communities, art could primarily function as a vehicle for re-establishing predictability. In Western culture, on the other hand, we artificially create obstacles for predictability (in art, disaster movies, etc) to be able to experience the joys of their resolution (while we often still use predictable patterns for our wallpapers and decorations). The proximate cause for making and consuming art would then be emotion regulation, not "just" uncertainty reduction. Also, according to Dissanayake, art can be traced back to the simplified, repeated and stereotyped interactions between adults and children, which assist the development of emotional self-regulation, attention and learning. Indeed, the dynamics of prediction and emotion seem to be protracted in children, where, for instance, in the in the peek-a-boo game the contrast effect (positive emotion following a negative one is more intense) is easily observed.

While art may therefore be a form of training of our exploratory learning capacities in a safe, playful context—information foraging is after all a vital human capacity (Vessel 2004)—it didn't necessarily evolve for that reason. Rather our approach connects to the neural recycling hypothesis (Dehaene and Cohen 2007), which assumes that art (similar

to, for example, writing) didn't evolve for any particular adaptive function but is the result of cultural inventions exploiting evolutionarily older brain circuits and inheriting many of their structural constraints. Indeed, there is no art module in the brain that needed to evolve. Artistic abilities are piggy-backing on our perceptual and emotional information processing capacities. Once in place, art may or may not have become a criterion by itself for selective forces to work on (co-optation), for example, in mate selection (Pinker 2003) or as a way to promote belonging to a social group.

Fourth, unpredictability and its resolution are important in other human activities. For instance, games are most rewarding when they have just the right amount of difficulty, of unpredictability. Similarly in humor we build up expectations and create discrepancies (Hurley et al 2011). Even in science we are most astonished when a scholar discovers (and manages to explain or make predictable) a counter-intuitive discrepancy in a very familiar domain. One might wonder what, if anything, is special about art. But does art need such an essence? Could we not suffice by stating that it is a human activity involving the full emotional and cognitive abilities of human beings but with no immediate biological purpose? A quasi-necessary result of a greatly expanded predictive capacity and an extended ability to delay gratification?

Finally, how do we explain within the predictive coding framework that humans—while ultimately aimed at maximizing predictability, or equivalently, minimizing prediction errors—still explore unpredictable stimuli, and even intentionally create them, as in the case of art. This is a matter of current debate (Fiorillo 2010; Friston 2010a, 2010b). The immediate motivation of seeking prediction errors may, in our view, be obtaining a larger reward (by contrast) later. Friston explains that this complicated exploratory (itinerant) behavior does not violate the general tenet of minimization of surprise, provided that the agent "revisits a small set of states, called a global random attractor that are compatible with survival" (Friston 2010a, p 2). His generalization of predictive coding in the free energy principle optimizes this motion through sensory state-space.

Turning to the advantages of our approach, artists and art critics will approve of the importance of the active, albeit largely implicit, role of the subject (the viewer) in the predictive coding approach. Even a static painting becomes a dynamic experience, as Dewey (2005, p 222) writes: "The product of art—temple, painting, statue, poem—is not the work of art. The work takes place when a human being cooperates with the product so that the outcome is an experience that is enjoyed because of its liberating and ordered properties." Our proposal also repeats another adage of artists: learn the rules well, so you can break them effectively.

Thus, predictive coding seems to agree with how artists themselves think about what they do. The painter Henri Matisse famously said when interviewed: "C'est une création par les rapports: Je ne peins pas les choses, je ne peins que des différences entre les choses" (Aragon 1971, p 140).[2] We may not read too much into it if we say that he intuitively appreciated that our brain works contextually, by figuring out the differences (prediction errors), not the absolute values (Nikolić 2010; Ramachandran and Blakeslee 1999). Any outline sketch is a sketch of the differences that the visual system would pick up when viewing the real scene. In fact, scene category can be decoded from fMRI activity in the visual system during the viewing of line drawings of scenes, just as well as from brain activity while viewing colour photographs of the scenes (Walther et al 2011).

In addition, honoring Occam's razor, we may not need a special, separate psychological theory for aesthetics by using predictive coding. This is consistent with the idea that

---

[2] "It's a creation by relation: I do not paint things, I only paint the differences between things" (Author's translation).

general-purpose motivations and capacities are involved in art, even though their particular combination might be special to art. In this case, we can rely on the cognitive and neural evidence for predictive coding, which is broader than visual perception (Friston 2010a; Winkler et al 2009). The latter is particularly appealing because art obviously consists of more than visual perception alone. Predicting is the default mode of the brain, encompassing perceptual and semantic levels. And even though the emotional implications of the predictive coding approach have not been thoroughly explored, its potential to connect perception, learning, and emotion may be clear from this proposal.

Lastly, our view may open new avenues for the empirical study of aesthetic appreciation. Prediction error and confirmation may be tractable in the lab, and thus may allow us to isolate one mechanism involved in aesthetic appreciation. For instance, we can induce strong short-term predictions in subjects and subsequently violate or confirm them. Also, it may help to have physiological markers of prediction violation, for example, in the event-related brain potential visual mismatch negativity (Kimura et al 2011). Here we can expect that temporal aspects and expertise will critically influence the outcomes in perception and emotion.

In this paper we have outlined a theory of art starting from the hierarchical, bidirectional dynamics of vision. We concluded that it is not the most predictable stimulus that is most pleasurable (the easiest Gestalt formation, cf perceptual fluency) but the Gestalt that appears unexpectedly after a fair amount of "obstinate obstruction". The positive affective evaluations result from a transition rather than a certain state of stimulation. A stimulus has to play hard to get before it can be pleasing (Lehrer 2008). But because our cognitive system ultimately aims to return to predictability, an optimal amount of prediction error exists. Eventually understanding art implies fully understanding our brain (not just the visual system) and its embodied embeddedness in the natural, social, and cultural environment. We are far from such an understanding, so any theorizing on art is necessarily preliminary and speculative. Our theory on art is really a theory about perception and emotion and their interplay because we believe that only by understanding this interaction we will come to comprehend human artistic behavior.

References

Alink A, Schwiedrzik C M, Kohler A, Singer W, Muckli L, 2010 "Stimulus Predictability Reduces Responses in Primary Visual Cortex" *The Journal of Neuroscience* **30** 2960–2966 ◄

Aragon L, 1971 *Henri Matisse: Roman* (Paris: Gallimard) ◄

Arnheim R, 1974 *Art and Visual Perception* (Berkeley, CA: University of California Press) ◄

Augustin M D, Wagemans J, Carbon C C, in press "All is beautiful? Generality vs. specificity of word usage in visual aesthetics" *Acta Psychologica* ◄

Bar M, 2007 "The proactive brain: Using analogies and associations to generate predictions" *Trends in Cognitive Sciences* **11** 280–289 doi:10.1016/j.tics.2007.05.005 ◄

Baron-Cohen S, Richler J, Bisaraya D, Gurunathan N, Wheelwright S, 2003 "The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences" *Philosophical Transactions of the Royal Society B: Biological Sciences* **358** 361–374 doi:10.1098/rstb.2002.1206 ◄

Barrett L F, Bar M, 2009 '"See it with feeling: affective predictions during object perception" *Philosophical Transactions of the Royal Society B: Biological Sciences* **364** 1325–1334 doi:10.1098/rstb.2008.0312 ◄

Baxter M G, Murray E A, 2002 "The amygdala and reward" *Nature Reviews Neuroscience* **3** 563–573 doi:10.1038/nrn875 ◄

Berlyne D E, 1970 "Novelty, complexity, and hedonic value" *Perception & Psychophysics* **8** 279–286 doi:10.3758/BF03212593 ◄

Biederman I, Vessel E, 2006 "Perceptual Pleasure and the Brain" *American Scientist* **94** 247–247 ◄

Blood A J, Zatorre R J, 2001 "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion" *Proceedings of the National Academy of Sciences* **98** 11818–11823 doi:10.1073/pnas.191355898 ◄

Bornstein R F, 1989 "Exposure and affect: Overview and meta-analysis of research, 1968–1987" *Psychological Bulletin* **106** 265–265 doi:10.1037/0033-2909.106.2.265 ◄

Bourdieu P, 1984 *Distinction: A Social Critique of the Judgement of Taste* (Cambridge, MA: Harvard University Press) ◄

Bromberg-Martin E S, Hikosaka O, 2011 "Lateral habenula neurons signal errors in the prediction of reward information" *Nature Neuroscience* **14** 1209–1216 doi:10.1038/nn.2902 ◄

Brown S, Gao X, Tisdelle L, Eickhoff S B, Liotti M, 2011 "Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities" *NeuroImage* **58** 250–258 doi:10.1016/j.neuroimage.2-011.06.012 ◄

Cerra P L, Bingham R, 1998 "The adaptive nature of the human neurocognitive architecture: An alternative model" *Proceedings of the National Academy of Sciences* **95** 11290–11294 doi:10.1073/pnas.95.19.11290 ◄

Clark A, 2011 "Predictive Coding" Edge World Question Center 2011, http://edge.org/q2011/q11_6.h-tml#clark ◄

Craig A D, 2009 "How do you feel—now? The anterior insula and human awareness" *Nature Reviews Neuroscience* **10** 59–70 doi:10.1038/nrn2555 ◄

Dehaene S, Cohen L, 2007 "Cultural Recycling of Cortical Maps" *Neuron* **56** 384–398 doi:10.1016/j.n-euron.2007.10.004 ◄

Dewey J, 2005 *Art as Experience* (New York: Penguin) ◄

de-Wit L, Machilsen B, Putzeys T, 2010 "Predictive Coding and the Neural Response to Predictable Stimuli" *Journal of Neuroscience* **30** 8702–8703 doi:10.1523/JNEUROSCI.2248-10.2010 ◄

Dissanayake E, 2008 "The arts after Darwin: Does art have an origin and adaptive function" *World art studies: Exploring concepts and approaches* (Amsterdam: Valiz) ◄

Dissanayake E, 2009 '"The Artification Hypothesis and Its Relevance to Cognitive Science, Evolutionary Aesthetics, and Neuroaesthetics" *Cognitive Semiotics* **9** 136–158 doi:10.3726/81609$_1$36 ◄

Dodgson N A, 2009 "Balancing the expected and the surprising in geometric patterns" *Computers & Graphics* **33** 475–483 doi:10.1016/j.cag.2009.04.001 ◄

Duncan S, Barrett L F, 2007 "Affect is a form of cognition: A neurobiological analysis" *Cognition & Emotion* **21** 1184–1211 doi:10.1080/02699930701437931 ◄

Dutton D G, Aron A P, 1974 "Some evidence for heightened sexual attraction under conditions of high anxiety" *Journal of Personality and Social Psychology* **30** 510–517 doi:10.1037/h0037031 ◄

Egner T, Monti J M, Summerfield C, 2010 "Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream" *Journal of Neuroscience* **30** 16601–16608 doi:10.1523/JNEUROSCI.2770-10.2010 ◄

Elliott R, Friston K, Dolan R J, 2000 "Dissociable Neural Responses in Human Reward Systems" *Journal of Neuroscience* **20** 6159–6165 ◄

Eysenck H J, 1942 "The experimental study of the 'good Gestalt'—a new approach" *Psychological Review* **49** 344–364 doi:10.1037/h0057013 ◄

Fairhall S L, Ishai A, 2008 "Neural correlates of object indeterminacy in art compositions" *Consciousness and Cognition* **17** 923–932 doi:10.1016/j.concog.2007.07.005 ◄

Fiorillo C D, 2010 "A neurocentric approach to Bayesian inference" *Nature Reviews Molecular Cell Biology* **11** 605–605 doi:10.1038/nrm2963 ◄

Friston K, 2010a "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience* **11** 127–138 doi:10.1038/nrn2787 ◄

Friston K, 2010b "Is the free-energy principle neurocentric?" *Nature Reviews Neuroscience* **11** 605–605 doi:10.1038/nrn2787 ◄

Gaver W, Mandler G, 1987 "Play it again, Sam: On Liking Music" *Cognition & Emotion* **1** 259–282 doi:10.1080/02699938708408051 ◄

Hebb D O, 1949 *The Organization of Behavior* (New York: John Wiley) ◄

Herry C, Bach D R, Esposito F, Di Salle F, Perrig W J, Scheffler K, Luthi A, Seifritz E, 2007 "Processing of Temporal Unpredictability in Human and Animal Amygdala" *Journal of Neuroscience* **27** 5958–5966 doi:10.1523/JNEUROSCI.5218-06.2007 ◄

Hurley M M, Dennett D C, Adams R B Jr, 2011 *Inside Jokes: Using Humor to Reverse-Engineer the Mind* (USA: MIT Press) ◄

Huron D, 2006 *Sweet Anticipation: Music and the Psychology of Expectation* (Cambridge, MA: MIT Press) ◄

Hyman J, 2010 "Art and Neuroscience" *Beyond Mimesis and Convention* R Frigg, M Hunter (Eds) (Dordrecht: Springer Netherlands) ◄

Ishai A, Fairhall S L, Pepperell R, 2007 "Perception, memory and aesthetics of indeterminate art" *Brain Research Bulletin* **73** 319–324 doi:10.1016/j.brainresbull.2007.04.009 ◄

Jackendoff R, Lerdahl F, 2006 "The capacity for music: What is it, and what's special about it?" *Cognition* **100** 33–72 doi:10.1016/j.cognition.2005.11.005 ◄

Johnson A W, Gallagher M, 2010 "Greater effort boosts the affective taste properties of food" Proceedings of the Royal Society B: Biological Sciences doi:10.1098/rspb.2010.1581 ◄

Kagan J, 2002 *Surprise, Uncertainty, and Mental Structures* (Cambridge, MA: Harvard University Press) ◄

Kapp B, Whalen P, Supple W, le , le , 1992 *"Amygdaloid contributions to conditioned arousal and sensory information processing" The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction, J Aggleton (Ed)* (New York: Wiley-Liss) ◄

Kimura M, Schröger E, Czigler I, 2011 "Visual mismatch negativity and its importance in visual cognitive sciences" *NeuroReport* **22** 669–673 doi:10.1097/WNR.0b013e32834973ba ◄

Koffka K, 1935 *"Principles of Gestalt psychology"* (New York: Harcourt-Brace) ◄

Kubovy M, 1999 *"On the Pleasures of the Mind" Well-being: The Foundations of Hedonic Psychology, D Kahneman, E Diener, N Schwarz (Eds)* (New York: Russell Sage Foundation) ◄

Kumar S, Sedley W, Nourski K V, Kawasaki H, Oya H, Patterson R D, Howard M A, Friston K J, Griffiths T D, 2011 "Predictive Coding and Pitch Processing in the Auditory Cortex" *Journal of Cognitive Neuroscience* **23** 3084–3094 doi:10.1162/jocn$_{a0}$0021 ◄

Landau M J, Greenberg J, Solomon S, Pyszczynski T, Martens A, 2006 "Windows Into Nothingness: Terror Management, Meaninglessness, and Negative Reactions to Modern Art" *Journal of Personality and Social Psychology* **90** 879–892 doi:10.1037/0022-3514.90.6.879 ◄

Lauwereyns J, 2010 *The Anatomy of Bias: How Neural Circuits Weigh the Options* (Cambridge MA: MIT Press) ◄

Leder H, Belke B, Oeberst A, Augustin D, 2004 "A model of aesthetic appreciation and aesthetic judgments" *British Journal of Psychology* **95** 489–508 doi:10.1348/0007126042369811 ◄

Lee T S, Mumford D, 2003 "Hierarchical Bayesian inference in the visual cortex" *Journal of the Optical Society of America A* **20** 1434–1448 doi:10.1364/JOSAA.20.001434 ◄

Lehrer J, 2008 *Proust Was a Neuroscientist* (New York: Houghton Mifflin Harcourt) ◄

Lindell A, Mueller J, 2011 "Can science account for taste? Psychological insights into art appreciation" *Journal of Cognitive Psychology* **23** 453–475 doi:10.1080/20445911.2011.539556 ◄

Ludmer R, Dudai Y, Rubin N, 2011 "Uncovering Camouflage: Amygdala Activation Predicts Long-Term Memory of Induced Perceptual Insight" *Neuron* **69** 1002–1014 doi:10.1016/j.neuron.2011.0-2.013 ◄

Mamassian P, 2008 "Ambiguities and conventions in the perception of visual art" *Vision Research* **48** 2143–2153 doi:10.1016/j.visres.2008.06.010 ◄

Mandler G, 2003 *"Emotion" History of Psychology Handbook of psychology, D K Freedheim (Ed)* (New York: John Wiley and Sons) ◄

McNally G P, Johansen J P, Blair H T, 2011 "Placing prediction into the fear circuit" *Trends in Neurosciences* **34** 283–292 doi:10.1016/j.tins.2011.03.005 ◄

Melcher D, Bacci F, 2008 "The visual system as a constraint on the survival and success of specific artworks" *Spatial Vision* **21** 347–362 doi:10.1163/156856808784532626 ◄

Meyer L B, 1961 *Emotion and Meaning in Music* (Chicago, IL: University of Chicago Press) ◄

Murray S O, Schrater P, Kersten D, 2004 "Perceptual grouping and the interactions between visual cortical areas" *Neural Networks: The Official Journal of the International Neural Network Society* **17** 695–705 ◄

Mueller J, Melwani S, Goncalo J A, in press "The Bias Against Creativity: Why People Desire But Reject Creative Ideas" *Psychological Science* ◄

Nikolić D, 2010 "The brain is a context machine" *Review of Psychology* **17** 33–38 ◄

Niv Y, Chan S, 2011 "On the value of information and other rewards" *Nature Neuroscience* **14** 1095–1097 doi:10.1038/nn.2918 ◄

Norton M I, Mochon D, Ariely D, in press "The IKEA effect: When labor leads to love" *Journal of Consumer Psychology* doi:10.1016/j.jcps.2011.08.002 ◄

Noyce A, Sekuler R, 2011 "Surprises are mistakes: An EEG source localization study of prediction errors" *Journal of Vision* **11** 182–182 doi:10.1167/11.11.182 ◄

Ogawa H, Watanabe K, 2011 "Implicit learning increases preference for predictive visual display" *Attention, Perception, & Psychophysics* **73** 1815–1822 doi:10.3758/s13414-010-0041-2

den Ouden H E M, Daunizeau J, Roiser J, Friston K J, Stephan K E, 2010 "Striatal Prediction Error Modulates Cortical Coupling" *Journal of Neuroscience* **30** 3210–3219 doi:10.1523/JNEUROSCI.44-58-09.2010 ◄

Pelowski M, Akiba F, 2011 "A model of art perception, evaluation and emotion in transformative aesthetic experience" *New Ideas in Psychology* **29** 80–97 doi:10.1016/j.newideapsych.2010.04.001 ◄

Pepperell R, 2011 "Connecting art and the brain: An artist's perspective on visual indeterminacy" *Frontiers in Human Neuroscience* **5** 84–84 doi:10.3389/fnhum.2011.00084 ◄

Pinker S, 2003 *The blank slate: The modern denial of human nature* (New York: Penguin) ◄

Pinna B, 2010 "New Gestalt principles of perceptual organization: an extension from grouping to shape and meaning" *Gestalt Theory* **32** 1–67 ◄

Preuschoff K, Quartz S R, Bossaerts P, 2008 "Human Insula Activation Reflects Risk Prediction Errors As Well As Risk" *Journal of Neuroscience* **28** 2745–2745 doi:10.1523/JNEUROSCI.4286-07.2008 ◄

Proulx T, Heine S J, Vohs K D, 2010 "When Is the Unfamiliar the Uncanny? Meaning Affirmation After Exposure to Absurdist Literature, Humor, and Art" *Personality and Social Psychology Bulletin* **36** 817–829 doi:10.1177/0146167210369896 ◄

Ramachandran V S, Blakeslee S, 1999 *Phantoms in the Brain: Probing the Mysteries of the Human Mind* (New York: Harper Perennial) ◄

Ramachandran V S, Hirstein W, 1999 "The science of art: a neurological theory of aesthetic experience" *Journal of Consciousness Studies* **37** 6–6 ◄

Rao R P N, Ballard D H, 1999 "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects" *Nature Neuroscience* **2** 79–87 doi:10.1038/4580 ◄

Reber R, Schwarz N, Winkielman P, 2004 "Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience?" *Personality and Social Psychology Review* **8** 364–382 doi:10.1207/s15327957pspr0804$_3$ ◄

Redies C, 2007 "A universal model of esthetic perception based on the sensory coding of natural stimuli" *Spatial Vision* **21** 97–117 doi:10.1163/156856807782753886 ◄

Sander D, Grafman J, Zalla T, 2003 "The human amygdala: an evolved system for relevance detection" *Reviews in the Neurosciences* **14** 303–316 doi:10.1515/REVNEURO.2003.14.4.303 ◄

Schultz W, Dayan P, Montague P R, 1997 "A Neural Substrate of Prediction and Reward" *Science* **275** 1593–1599 doi:10.1126/science.275.5306.1593 ◄

Sloboda J A, 1991 "Music Structure and Emotional Response: Some Empirical Findings" *Psychology of Music* **19** 110–110 doi:10.1177/0305735691192002 ◄

Sperber D, 2005 *"Modularity and relevance: How can a massively modular mind be flexible and context-sensitive?" The Innate Mind: Structure and Content* (New York: Oxford University Press) ◄

Sperber D, Hirschfeld L A, 2004 "The cognitive foundations of cultural stability and diversity" *Trends in Cognitive Sciences* **8** 40–46 doi:10.1016/j.tics.2003.11.002 ◄

Spinoza B de, 2001 *Ethics* (London: Wordsworth Editions) ◄

Summerfield C, Monti J M P, Trittschuh E H, Mesulam M-M, Egner T, 2008 "Neural repetition suppression reflects fulfilled perceptual expectations" *Nature Neuroscience* **11** 1004–1006 doi:10.1038/nn.2163 ◄

Thompson M M, Naccarato M E, Parker K C H, Moskowitz G B, 2001 "The personal need for structure and personal fear of invalidity measures: Historical perspectives, current applications, and future directions" *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* ◄

Topolinski S, Reber R, 2010 "Gaining Insight Into the "Aha" Experience" *Current Directions in Psychological Science* **19** 402–402 doi:10.1177/0963721410388803 ◄

Van de Cruys S, Wagemans J, in press "Gestalts as predictions" *Gestalt Theory* ◄

van Leeuwen C, 2007 "What needs to emerge to make you conscious?" *Journal of Consciousness Studies* **14** 115–136 ◄

Verstijnen I M, van Leeuwen C, Goldschmidt G, Hamel R, Hennessey J M, 1998 "Sketching and Creative Discovery" *Design Studies* **19** 519–546 doi:10.1016/S0142-694X(98)00017-9 ◄

Vessel E A, 2004 "Behavioral and Neural Investigation of Perceptual Affect" Unpublished PhD thesis University of Southern California ◄

Vygotsky L S, 1971 *Psychology of Art* (Cambridge, MA: MIT Press) ◄

Walther D B, Chai B, Caddigan E, Beck D M, Fei-Fei L, 2011 "Simple line drawings suffice for functional MRI decoding of natural scene categories" *Proceedings of the National Academy of Sciences* **108** 9661–9666 doi:10.1073/pnas.1015666108 ◄

Whalen P J, 1998 "Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala" *Current Directions in Psychological Science* **7** 177–188 doi:10.1111/1467-8721.ep10836912 ◄

Winkler I, Denham S L, Nelken I, 2009 "Modeling the auditory scene: predictive regularity representations and perceptual objects" *Trends in Cognitive Sciences* **13** 532–540 doi:10.1016/j.tics.2009.09.003 ◄

Winstanley C A, Cocker P J, Rogers R D, 2011 "Dopamine Modulates Reward Expectancy During Performance of a Slot Machine Task in Rats: Evidence for a 'Near-miss' Effect" *Neuropsychopharmacology* **36** 913–925 doi:10.1038/npp.2010.230 ◄

Witkin H A, Goodenough D R, 1981 "Cognitive styles: essence and origins. Field dependence and field independence" *Psychological Issues* **51** 1–141 ◄

Yarkoni T, Poldrack R A, Nichols T E, Van Essen D C, Wager T D, 2011 "Large-scale automated synthesis of human functional neuroimaging data" *Nature Methods* **8** 665–670 doi:10.1038/nmeth.1635 ◄

Zajonc R B, 1968 "Attitudinal Effects of Mere Exposure" *Journal of Personality and Social Psychology* **9** 1–27 doi:10.1037/h0025848 ◄

Zeki S, 2004 "The neurology of ambiguity" *Consciousness and Cognition* **13** 173–196 doi:10.1016/j.concog.2003.10.003 ◄

Sander Van de Cruys (1982) has a MA in social and cultural anthropology and in psychology from the University of Leuven, where he is currently pursuing a PhD in experimental psychology. In his research he focuses on top- down influences in visual perception and on the interplay between perception and emotion (see www.gestaltrevision.be).

Johan Wagemans (1963) has a BA in psychology and philosophy, an MSc and a PhD in experimental psychology, all from the University of Leuven, where he is currently a full professor. Current research interests are mainly in so-called mid-level vision (perceptual grouping, figure-ground organization, depth and shape perception) but stretching out to low-level vision (contrast detection and discrimination) and high-level vision (object recognition and categorization), including applications in autism, arts, and sports (see http://www.gestaltrevision.be).